Purification and Characterization of Tryptophan Hydroxylase

Haahr, Lærke Tvedebrink

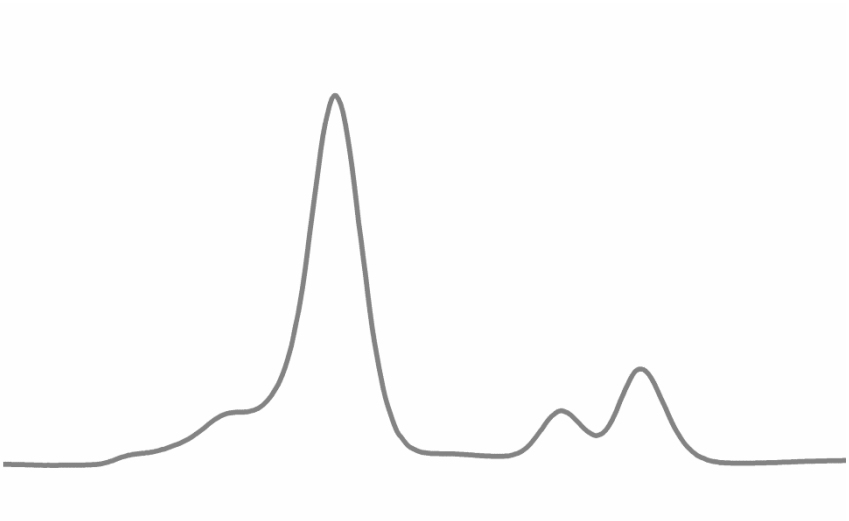
Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Technical University of Denmark

Purification and Characterization of Tryptophan Hydroxylase

Lærke Tvedebrink Haahr



PhD Thesis
Metalloprotein Chemistry & Engineering
DTU Chemistry
Technical University of Denmark
2012

Purification and Characterization of Tryptophan Hydroxylase

Lærke Tvedebrink Haahr

Supervisor:
Associate Professor Hans Erik Mølager Christensen

PhD Thesis
Metalloprotein Chemistry & Engineering
DTU Chemistry
Technical University of Denmark
2012

Front-page picture: Gel-filtration chromatogram

Preface and acknowledgement

The present dissertation entitled “Purification and Characterization of Tryptophan Hydroxylase” is submitted to the Technical University of Denmark and partially fulfils the requirements for the PhD degree. The work has been carried out at DTU Chemistry in the Metalloprotein Chemistry & Engineering group with Associate Professor Hans E. M. Christensen as supervisor. The PhD scholarship was funded by the Technical University of Denmark.

First and foremost, I would like to thank my supervisor Associate Professor Hans E. M. Christensen for his help and guidance during this project and for always being available for questions and inspiring scientific discussions. I also thank him for encouraging me when nothing worked out and keeping me realistic when I got a little too excited.

I would like to thank the entire Metalloprotein Chemistry & Engineering group, both past and present, for providing a pleasant working atmosphere. A special thanks to laboratory technician Lise-Lotte Jespersen for being a great help in the laboratory, and for taking good care of us all. Also thanks to Jane Boesen and Michael S. Windahl for their previous work on tryptophan hydroxylase.

During this project I have co-supervised several bachelor students, and I would like to thank them for their contributions to the experimental work.

I would also like to thank Professor Edward I. Solomon, Department of Chemistry, Stanford University for allowing me to join his group and carry out circular dichroism experiments. Thanks to graduate student Lei V. Liu from the Solomon Laboratory for his help with measurements and subsequent data analysis. I would like to thank the Ministry of Science Technology and Innovation for awarding me the EliteForsk travel-scholarship, which provided financial support for my stay in USA along with several conferences and synchrotron trips.

Finally, I would like to thank Associate Professor Ole Kristensen and Associate Professor Lars Olsen, both from Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, Copenhagen University for their help with crystallization and isothermal titration calorimetry, respectively. Also thanks to Facility Manager Dr. Maria Håkansson for her help with differential scanning fluorimetry.

_____

Lærke Tvedebrink Haahr
Kgs. Lyngby, June 2012

Dansk resumé

Denne afhandling omhandler oprensning og karakterisering af jern enzymet tryptophanhydroxylase (TPH). TPH eksisterer i to isoformer, TPH1 og TPH2. Hver isoform består af 3 strukturelt forskellige domæner: det regulatoriske, det katalytiske og det tetrameriserende domæne. TPH katalyserer omdannelsen af tryptophan til 5-hydroxytryptophan, hvilket er det hastighedsbestemmende trin i biosyntesen af serotonin. Serotonin er en vigtig neurotransmitter, som er involveret i en række psykiske sygdommme såsom depression og tvangstanker. Målet med dette projekt var at udvikle oprensningsmetoder for fuldlængde TPH1 og TPH2 samt at karakterisere de oprensede TPH varianter.

Der blev udviklet en succesfuld oprensningsmetode for fuldlængde human TPH1 (*h*TPH1), som resulterede i rent, aktivt og stabilt protein. Metoden inkluderede affinitetsoprensning med maltose bindende protein (MBP) tag og høje salt koncentrationer i alle buffere. Det er første gang en succesfuld oprensningsmetode for *h*TPH1 præsenteres, som også opnår stabilt protein efter kløvning fra affinitetstagget. Ydermere blev der udviklet succesfulde oprensningsprocedurer for en række andre TPH varianter. En oprensningsmetode for det regulatoriske og katalytiske domæne af TPH1 (rc*h*TPH1) med MBP tag blev udviklet, og afslørede at det regulatoriske domæne forårsager dimerisering. En ny oprensningsmetode for det katalytiske domæne at human TPH1 (c*h*TPH1) med gluthathione S-transferase (GST) tag blev også udviklet. To humane TPH2 varianter indeholdende det regulatoriske domæne blev forsøgt oprenset efter den samme generelle metode som *h*TPH1, men udbytterne var lave og produkterne ustabile og kunne ikke opkoncentreres. Den høje salt koncentration stabiliserer altså ikke det regulatoriske domæne af TPH2. Stabiliteten af 4 TPH varianter blev undersøgt vha. differential skanningsfluorimetri (DSF) and resultaterne giver information om effekten af pH, salt type og ionstyrke.

Resultaterne fra krystallisationsarbejdet var desværre meget begrænsede. Krystallisationsproceduren for the katalytiske domæne af *gallus gallus* TPH1 (c*g*TPH1) blev optimeret til hurtigere krystal dannelse ved tilsætning af tryptophan og inkubation ved stuetemperatur. Krystaller uden imidazole i krystallisationsbetingelserne blev lavet. De løste strukturer var dog af dårlig kvalitet og bidrager ikke med ny viden om TPH. En strategisk fremgangsmåde til krystallisation af TPH varianter blev udviklet, og screeningen producerede lovende betingelser for rc*h*TPH1 og flere krystaller for c*h*TPH1.

Fire mutanter af det katalytiske domæne af human TPH2 (c*h*TPH2) blev fremstillet, og det kan konkluderes ar Glu363, som er jern ligand, er essentiel for aktivitet, mens Tyr358 påvirker aktiviteten, uden at være essentiel. Cirkulærdikroisme (CD) og magnetisk cirkulærdikroisme (MCD) resultater viste at jern i det aktive site skifter konformation fra 6 til 5 koordineret kun når både substrat og cofaktor binder. Der blev observeret forskelle mellem prøver lavet i glycerol og sucrose. Tilsætning af glycerol til c*h*TPH2 med bundet Fe(III) resulterede i en gul farve som indikerer interaktion mellem de to. DSF målinger demonstrerede at c*h*TPH2 er signifikant mere stabil i Fe(II) formen end i Fe(III) formen. Ydermere stabiliserer tilsætningen af substrat og cofaktor kun c*h*TPH2 på Fe(II) formen. Isotermisk titreringskalorimetri målinger blev udført på c*h*TPH2 og dissociationskonstanten for BH_2 blev bestemt til 5 µM, men dissociationskonstanten for tryptophan var noget højere. Tryptophan og BH_2 binder uafhængigt af hinanden.

Generelt har arbejdet udført i dette projekt bidraget signifikant til forskningen i TPH, både ved at udvikle nye oprensningsprocedurer samt tilføje viden om mekanismen af TPH.

Abstract

This thesis deals with the purification and characterization of the iron-containing enzyme tryptophan hydroxylase (TPH). TPH exists in two isoforms, called TPH1 and TPH2. Each isoform consists of three structural distinct domains: the regulatory, the catalytic and the tetramerization domain. TPH catalyzes the hydroxylation of tryptophan to 5-hydroxytryptophan, which is the rate-limiting step in the biosynthesis of serotonin. Serotonin is an important neurotransmitter, which is involved in a range of psychiatric disorders including depression and obsessive-compulsive disorder. The goal of this project was to develop purification methods for full-length TPH1 and TPH2 as well as to characterize purified TPH variants.

A successful purification method for full-length human TPH1 (*h*TPH1) was developed, which resulted in pure, active and stable protein. The method includes affinity-purification using maltose binding protein (MBP) tag and high salt concentration in all buffers. It is the first time a successful purification method for *h*TPH1 has been presented, which also yields stable protein after cleavage from the affinity tag. Furthermore, successful purification procedures were developed for a number of other TPH variants. A purification method for the regulatory and catalytic domains of human TPH1 (rc*h*TPH1) using MBP affinity tag was developed, and revealed that the regulatory domain causes dimerization. Also, a new purification method for the catalytic domain of human TPH1 (c*h*TPH1) using gluthathione S-tranferase (GST) tag was developed. Two human TPH2 variants containing the regulatory domain were attempted purified using the same general method as for *h*TPH1, but the yields were very low and the products were unstable and could not be concentrated. The high salt concentration does thus not seem to stabilize the regulatory domain of TPH2. The thermal stability of 4 TPH variants was studied by differential scanning fluorimetry (DSF), and the results gave information on the effect of pH, salt type and ionic strength.

The results from crystallization were unfortunately rather limited. The crystallization procedure for the catalytic domain of *gallus gallus* TPH1 (c*g*TPH1) was optimized to faster crystal growth by addition of tryptophan and incubation at room temperature. Crystals without imidazole in the crystallization conditions could be obtained. The solved structures were however of poor quality and did not contribute with any new information on TPH. A strategic approach to crystallization of TPH variants was developed, and the pipeline screening produced leads for rc*h*TPH1 and several crystals for c*h*TPH1.

Four mutants of the catalytic domain of human TPH2 (c*h*TPH2) were made, and it can be concluded that Glu363, which is an iron ligand, is essential for activity, while Tyr358 affects the activity, but is not absolutely required. Circular dichroism and magnetic circular dichroism results on c*h*TPH2 clearly demonstrated that the active site iron changes conformation from 6 to 5 coordinated only when both substrate and cofactor binds. A difference between samples made in glycerol and sucrose was observed. Addition of glycerol to c*h*TPH2 in the Fe(III) form gave a yellow color indicative of interaction. DSF measurements demonstrated that c*h*TPH2 is significantly more stable in the Fe(II) form compared to the Fe(III) form. Furthermore, the addition of substrates and cofactor only stabilizes c*h*TPH2 when Fe(II) is present. Isothermal titration calorimetry measurements on c*h*TPH2 determined the dissociation constant of BH_2 to 5 µM, while the dissociation constant for tryptophan is somewhat higher. Tryptophan and BH_2 bind independently of each other.

Overall, the work performed in this project has made significant contributions to the research on TPH, both by providing new purification procedures as well as advancements in the knowledge on the mechanism of TPH.

List of abbreviations

100amp	100 µg/mL ampenicillin
30kan	30 µg/mL kanamycin sulfate
3CP	Human rhinovirus 3C protease
4a-OH-BH_4	4a-hydroxy-tetrahydrobiopterin
5C	5 coordinated
5-OH-trp	5-hydroxytryptophan
6C	6 coordinated
AAAH	Aromatic amino acid hydroxylases
ADHD	Attention deficit hyperactivity disorder
BH_2	7,8-dihydrobiopterin
BH_4	5,6,7,8-tetrahydrobiopterin
CD	Circular dichroism
CNS	Central nervous system
CV	Column volume
DOPA	3,4-dihydroxyphenylalanine
DSF	Differential scanning fluorimetry
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
GAPD	Glyceraldehyde-3-phosphate dehydrogenase
GST	Gluthathione S-transferase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High-performance liquid chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
LCP	Left circularly polarized
LF	Ligand field
MBP	Maltose binding protein
MCD	Magnetic circular dirchroism
oa-TOF	Orthogonal acceleration time-of-flight
OCD	Obsessive-compulsive disorder
OD_{600}	Optical density at 600 nm
PAH	Phenylalanine hydroxylase
PEG	Polyethylene glycol
PKU	Phenylketonuria
RCP	Right circularly polarized
RMS	Root mean square
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SOD	Superoxide dismutase
TCEP	Tris(2-carboxyethyl)phosphine
TH	Tyrosine hydroxylase
TPH	Tryptophan hydroxylase
TPH1	Tryptophan hydroxylase isoform 1
TPH2	Tryptophan hydroxylase isoform 2
trp	Tryptophan
UV-Vis	Ultraviolet-visible spectroscopy

Contents

1 Introduction

This thesis deals with the purification and characterization of the iron-containing enzyme tryptophan hydroxylase (TPH). TPH catalyzes the hydroxylation of tryptophan to 5-hydroxytryptophan, as seen in Figure 1, using tetrahydrobiopterin and dioxygen.

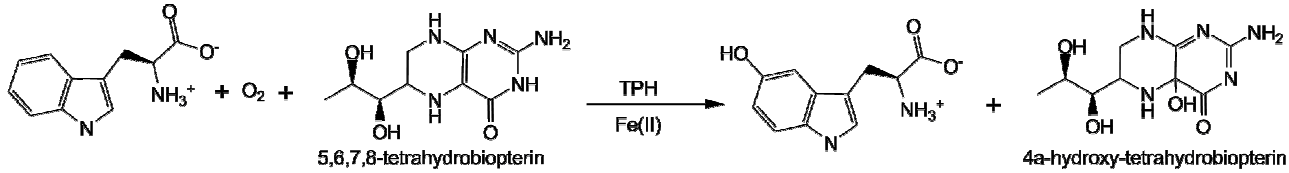


Figure 1: The reaction catalyzed by tryptophan hydroxylase (TPH). Tryptophan is hydroxylated to 5-hydroxytryptophan, while tetrahydrobiopterin is hydroxylated to 4a-hydroxy-tetrahydrobiopterin. O_2 is the oxygen source.

The reaction catalyzed by TPH is the first and rate-limiting step in the biosynthesis of serotonin(*1*). Serotonin is an important neurotransmitter and furthermore a precursor for the hormone melatonin(*2*), which is involved in the circadian rhythm control of the sleep-wake cycle(*3*). The biosynthesis of serotonin and melatonin is seen in Figure 2.

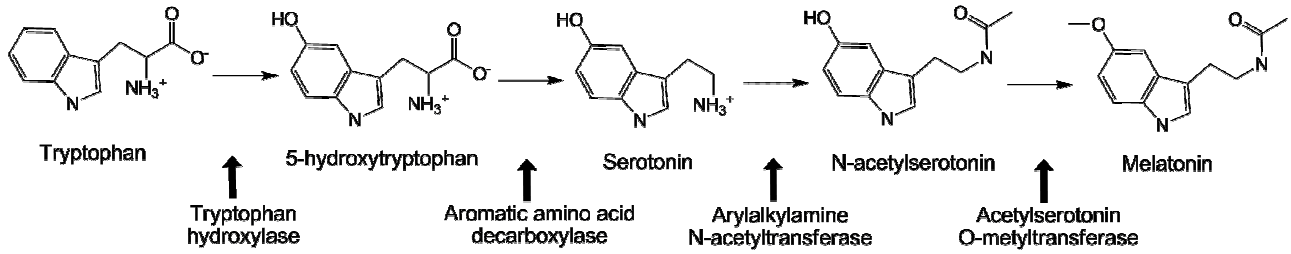


Figure 2: Overview of the biosynthesis of the neurotransmitter serotonin and the hormone melatonin from the amino acid tryptophan. Tryptophan hydroxylase catalyzes the reaction from tryptophan to 5-hydroxytryptophan.

Serotonin is involved in several physiological functions such as appetite and regulation of body temperature(*4*), as well as a range of psychiatric disorders including depression and obsessive-compulsive disorder(*5*). The importance of TPH in the production of serotonin has lead to a great interest in the enzyme, and large efforts have been made to characterize TPH in order to understand the function and regulation. Despite these efforts, knowledge on TPH is limited due to the inherent instability of the enzyme. In this thesis, attempts will be made to increase the information available on TPH both by providing effective purification procedures as well as by characterizing the purified TPH variants.

TPH exists in two isoforms, called TPH1 and TPH2(*6*). TPH1 is expressed in peripheral parts of the body, while TPH2 primarily is expressed in the central nervous system(*7*). Each isoform consists of three structural distinct domains: the N-terminal regulatory domain, the catalytic domain containing the active site and the small C-terminal tetramerization domain(*2*)(*8*). In the Metalloprotein Chemistry & Engineering Group, TPH has been studied for several years. As full-length TPH was found difficult to express and purify, truncated

variants of the different domains were made. Purification methods for truncated variants containing the catalytic domains of *gallus gallus* (chicken) TPH1, human TPH1 and human TPH2 have been developed(*9*, *10*). Also, a variant containing the catalytic and tetramerization domains of human TPH2 has been purified(*9*, *11*). These variants have all been kinetically characterized and some have also been structurally characterized(*9*–*13*). Recently, a purification method for the regulatory domain of human TPH1 was developed using affinity tags(*10*).

This project had two main goals: I) To develop purification methods for full-length human TPH1 and TPH2, and II) to characterize purified variants of TPH to achieve a better understanding of the structure-functional relationship. Large efforts were put into crystallization of TPH variants to solve the three-dimensional structures; however the results from this work were unfortunately very limited.

1.1 Outline of thesis

This thesis has been divided into three main chapters, concerning purification, crystallization and mechanistic studies, respectively. First, an introduction to the current relevant knowledge on TPH is given as an overview of the literature on the topic in chapter 2. In chapter 3, the background for and the strategies followed in this project are briefly presented. Chapter 4 concerns the development of suitable purification procedures for both TPH1 and TPH2 variants. The end-products from the purifications are in many cases characterized by mass spectrometry and the stability is studied by differential scanning fluorimetry. The attempts to structurally characterize the purified TPH variants by crystallization are described in chapter 5. Several different characterization methods were used in the mechanistic studies presented in chapter 6. The results from these are presented and discussed. Finally, in chapter 7 the overall conclusions from the work done in this project are given.

2 Tryptophan hydroxylase

In this chapter, a brief introduction to the relevant knowledge on tryptophan hydroxylase will be given. The information given here only represents an overview, and more specific and thorough discussions on relevant literature will be given in the following chapters according to topic.

2.1 The aromatic amino acid hydroxylases

Tryptophan hydroxylase (EC 1.14.16.4) is a non-heme iron-containing enzyme, which catalyzes the reaction between L-tryptophan (trp), O_2, and 5,6,7,8-tetrahydrobiopterin (BH_4) to give 5-hydroxytryptophan (5-OH-trp) and 4a-hydroxy-tetrahydrobiopterin (4a-OH-BH_4) (*14*).

BH_4 acts as electron donor in the reduction of O_2 and hydroxylation of trp. *In vivo* 4a-OH-BH_4 is regenerated by the enzymes pterin-4a-carbinolamin dehydratase and dihydrobiopteridine reductase producing H_2O at the expense of nicotinamide adenine dinucleotide (NADH)(*15*). Thus, TPH is defined as a mono-oxygenase and BH_4 as a cofactor, even though BH_4 actually acts as a co-substrate in the TPH-catalyzed reaction. For the rest of this thesis, BH_4 will be referred to as a cofactor, and trp and O_2 as substrates.

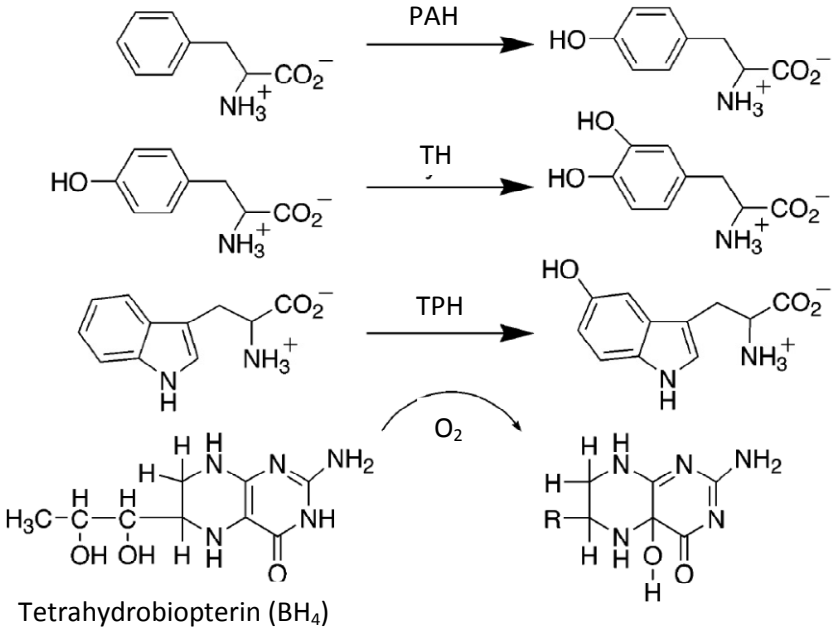


Figure 3: The reactions catalyzed by the aromatic amino acid hydroxylases, where the amino acid side chain is hydroxylated utilizing tetrahydrobiopterin as electron donor and O_2 as the oxygen source. PAH: Phenylalanine hydroxylase. TH: Tyrosine hydroxylase. TPH: Tryptophan hydroxylase. Adapted with permission from (*14*)**. Copyright 2003 American Chemical Society.**

Together with phenylalanine hydroxylase (PAH) and tyrosine hydroxylase (TH), TPH constitutes the enzyme family of aromatic amino acid hydroxylases (AAAHs)(*16*, *17*). The AAAHs all require Fe(II) and BH_4 for hydroxylation of their respective amino acids (see Figure 3), and show substantial similarities in their

structure and catalytic mechanism(*16*). The mammalian AAAHs form homotetramers, and each monomer consists of three distinct structural and functional domains: The N-terminal regulatory domain, the catalytic domain including the iron active site and substrate binding pockets, and the small C-terminal tetramerization domain(*14*).

PAH is found in the liver(*18*), where it catalyzes the rate-limiting step in catabolism of phenylalanine by converting phenylalanine to tyrosine (*19*). Deficiency in PAH lead to phenylketonuria (PKU), a disease where excess phenylalanine is converted to phenylpyruvate(*17*). Phenylpyrovate is toxic and can lead to mental retardation, but usually this can be avoided through a low phenylalanine-diet(*20*). In Denmark, all newborns are tested for PKU 48-72 hours after birth, to ensure early treatment. Only 1 in 11000 newborns is diagnosed with PKU in Denmark(*21*).

TH is found in the central nervous system (CNS) and adrenal gland, where it catalyzes the formation of 3,4-dihydroxyphenylalanine (DOPA) from tyrosine(*22*). This is the first and rate-limiting step in the biosynthesis of cathecolamine neurotransmitters such as dopamine, norepinephrine (noradrenaline), and epinephrine (adrenaline)(*20*). Deficiencies in TH have been related to the genetic disease L-DOPA responsive dystonia(*23*) and Parkinson's disease(*24*, *25*).

The sequential and structural homology of the catalytic domains of the AAAHs and the similarity in the reactions they catalyze, suggests that they share a common catalytic mechanism(*14*, *26*). Thus, knowledge on PAH and TH can be used for understanding the mechanism of TPH. The latter has been characterized to much less extent compared to PAH and TH due to the inherent instability of the protein, and hence the difficulty to purify it(*5*). The three AAAHs regulatory domains differ significantly in sequence and length, and different regulatory properties are observed(*14*).

2.2 The two isoforms of tryptophan hydroxylase

TPH exists in two isoforms(*6*), called TPH isoform 1 (TPH1) and TPH isoform 2 (TPH2). The two TPH isoforms are highly homologous with a sequence identity of 71 %(*7*), but are expressed in different tissues of the body. TPH1 appears to be mainly expressed in non-neuronal serotonergic cells, e.g. skin, pineal gland, and enterochromaffin cells in the intestines (*1*, *7*, *27–30*). TPH2 is expressed in neuronal serotonergic cells, e.g. serotonergic neurons of the brain and gut (*6*, *28*, *29*, *31*, *32*). TPH2 is thus responsible for synthesis of serotonin in the CNS, while TPH1 controls serotonin synthesis in the peripheral system.

The main differences between TPH1 and TPH2 are found in the regulatory domain, where the N-terminal is extended with 46 residues in TPH2(*7*), as shown in Figure 4. The extra amino acids in TPH2 are believed to reduce the expression level compared to TPH1(*33*), and to inhibit enzyme activity(*34*).

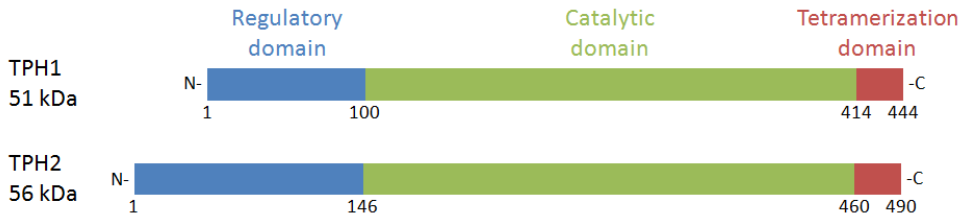


Figure 4: The division of the three domains in human TPH1 and human TPH2. The regulatory domains are shown in blue, the catalytic domains in green, and the tetramerization domains in red.

The peripheral and central serotonergic systems, which are regulated by TPH1 and TPH2, respectively, are coupled to different physiological functions and thus different disorders(*7*). TPH1 is involved in

gastrointestinal disorders associated with excess serotonin, and inhibitors for TPH1 shows positive results for treatment of chronic diarrhea(*35*), irritable bowel syndrome(*36*), and colitis(*37*). TPH1 is thus a drug target for gastro-intestinal disorders related to serotonin dysregulation. Dysfunction of the central serotonergic system is associated with several neuropsychiatric disorders, such as obsessive-compulsive disorder (OCD), anxiety disorders, attention deficit hyperactivity disorder (ADHD), bipolar affective disorder, anorexia, schizophrenia and depression(*4*, *38*, *39*). Since TPH2 catalyzes the rate-limiting step in the biosynthesis of serotonin in the CNS, a possible correlation between genetic variation in TPH2 and the occurrence of the above mentioned disorders have been extensively studied. Such correlations have indeed been found for depression(*40*), suicidal behavior(*41*), bipolar disorder(*42*), OCD(*43*), and ADHD(*44*). However, most of the association studies have failed to be confirmed in samples from others, and thus the exact role of TPH2 in neuropsychiatric disorders remain unclear(*5*, *45*). Despite this, TPH2 is still considered a potential drug target for serotonin related psychiatric disorders(*45*, *46*).

2.3 Structure

Only a limited number of TPH structures are available in the literature(*12*, *47*, *48*). Thus, some of the structural information comes from the related enzymes TH and PAH. The domains are presented here one by one.

2.3.1 Regulatory domain

The N-terminal regulatory domain is the most diverse domain among the AAAHs. No structure of the regulatory domain of TPH has been solved, but the structure of the regulatory and catalytic domain of PAH has been determined(*49*). In this structure, the first 18 amino acids are lacking, as there is no visible electron density. The N-terminal residues 19-33 extend across the catalytic domain restricting access to the active site, as seen in Figure 5. Movement of this part of the regulatory domain is believed to be involved in regulation of PAH(*49*).

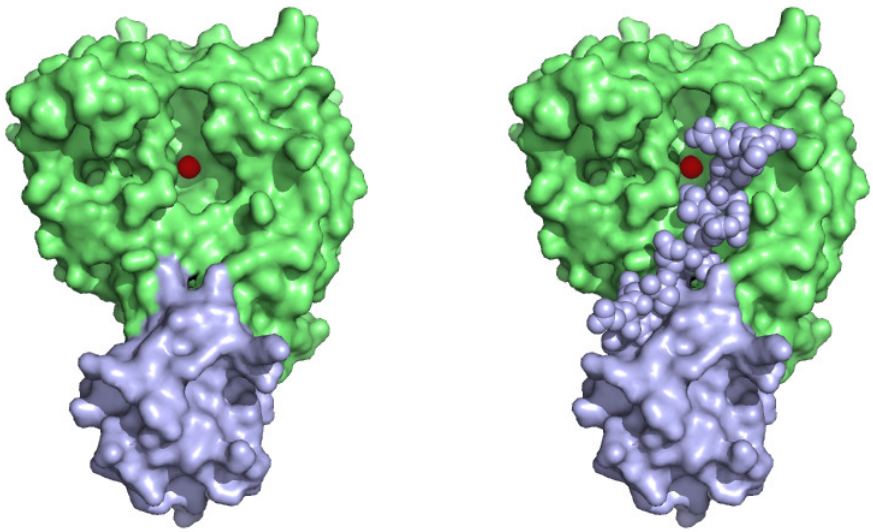


Figure 5: Structure of the regulatory (light blue) and catalytic (green) domains of PAH (PBD ID: 1PHZ). Iron is shown in red. Left: Residues 19-33 are not shown in the figure (residue 1-18 are lacking in the structure). The active site is open. Right: Residues 19-33 are included and shown as spheres. They clearly restrict the access to the active site. The figure was made using PyMOL(*50*)**.**

The regulation of TPH1 and TPH2 is not well understood(*5*). However, both isoforms are suggested to be regulated by phosphorylation by protein kinase A and calmodulin-dependent kinase(*51–54*) and interactions with 14-3-3 protein(*53*, *55*, *56*). The regulation of TPH expression is also being investigated(*33*, *57*).

2.3.2 Catalytic domain

5 structures of the catalytic domain of TPH1 are available from the PDB database. 4 of these are of the catalytic domain of human TPH1, one binding the cofactor analogue and inhibitor 7,8-dihydrobiopterin (BH_2)(*47*), and the 3 others binding designed inhibitors(*48*). A structure of the catalytic domain of *gallus gallus* TPH1 binding the substrate tryptophan was solved by the Metalloprotein Chemistry & Engineering group(*12*). The structures of the catalytic domain of TPH are very similar to those of PAH and TH as expected from the high sequence identity(*47*). The active site consists of a 9 Å deep and 10 Å wide cavity, with iron at the floor of this cavity 13 Å below the surface(*47*). Iron is octahedrally coordinated by His272, His277, Glu317 (human TPH1 numbering) and three water molecules in a so-called facial triad common to many non-heme iron oxygen activating enzymes(*58*). Iron is placed right at the intersection between the cofactor and substrate binding pockets, as seen in Figure 6.

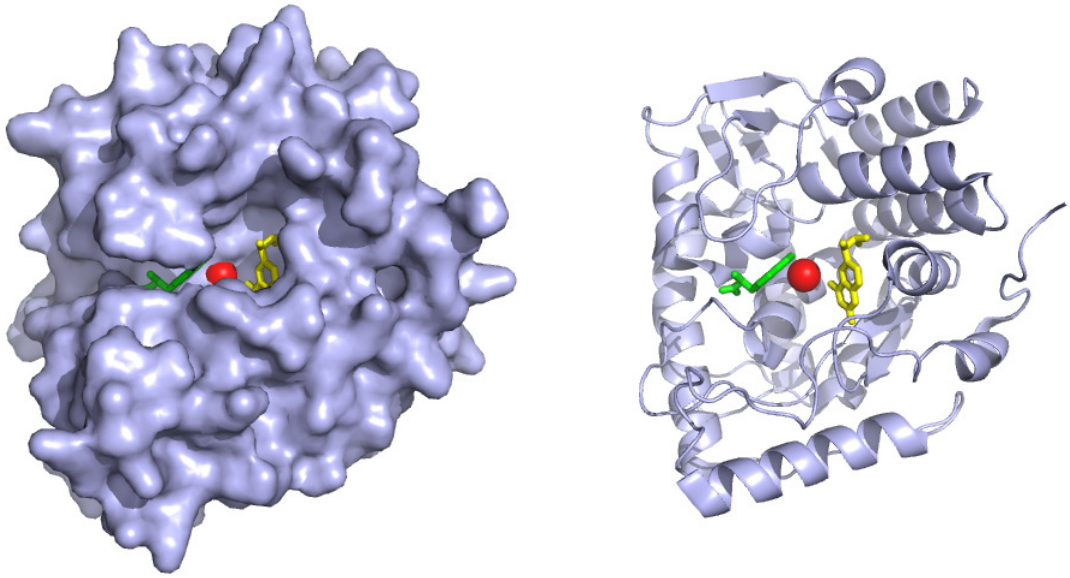


Figure 6: Structure of the catalytic domain of human TPH1 with BH_2 bound (PDB ID: 1MLW)(*47*). Tryptophan was added to the structure by alignment of the 3E2T structure using PyMOL. Iron is shown in red, tryptophan in green and BH_2 in yellow. Left: Surface representation showing the active site and the binding pockets for BH_2 and tryptophan. Right: Cartoon representation showing the overall fold. The figure was made using PyMOL(*50*).

2.3.2.1 Substrate binding site

From the structure of *gallus gallus* TPH1, the binding site of the natural substrate tryptophan can be investigated(*12*). A schematic overview of the interactions is shown in Figure 7. Tryptophan binds through polar interactions between the amino moiety and Arg258, Thr266, and Ser337, as well as between the side chain and Ile367. The side chain of tryptophan is bound in a hydrophobic pocket. The tryptophan stacks against Pro269 and there is a distance of 3.9 Å between the iron and the tryptophan C- atom that is hydroxylated (marked with * in Figure 7). Phe314 (*gallus gallus* TPH1 numbering) has an important role in defining substrate specificity. In PAH and TH the corresponding amino acid is tryptophan, which decrease the size of the hydrophobic binding pocket affecting substrate specificity. This has been investigated by

mutation of this Phe in TPH1 to Trp and the corresponding Trp in PAH to Phe, which significantly changed the substrate specificity (*59–61*). Also Tyr236 (*gallus gallus* TPH1 numbering), which is not conserved in PAH and TH, has been identified as important for substrate binding(*62*).

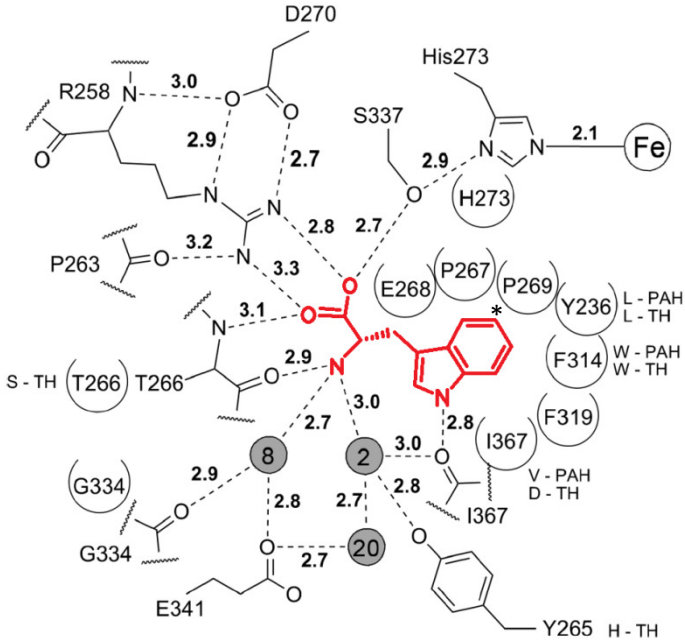


Figure 7: Schematic view of the tryptophan substrate interactions with *gallus gallus* TPH1 (PDB ID: 3E2T). Tryptophan is shown in red, hydrogen bonds as dashed lines, and the filled gray circles represent water molecules. Semi circled residues have hydrophobic interactions with the tryptophan substrate. For residues which are not conserved in PAH and TH, the corresponding residue is shown. The C-atom that is hydroxylated is marked with *. Adapted with permission from (*12*)**. Copyright 2008 American Chemical Society.**

2.3.2.2 Cofactor binding site

The cofactor binding site can be investigated in the structure of human TPH1 with the cofactor analogue BH_2 bound(*47*). The cofactor binds in the same orientation and position as in PAH. The pterin π-stacks on Phe241 and Tyr325 (human TPH1 numbering) and forms hydrogen bonds to Gly234 and Leu236. The O4 atom in the pterin is hydrogen bonded to two of the three water molecules coordinated to iron. Glu273 also forms two water-mediated hydrogen bonds to the pterin. The distance between the C-atom that is hydroxylated in pterin and iron is 5.7 Å.

2.3.3 Tetramerization domain

There is no structure of the tetramerization domain of TPH available in the literature. However, structures of the tetrameric forms of both TH(*63*) and PAH(*64*) have been solved. The small tetramerization domain at the C-terminal of the AAAHs is responsible for oligomerization through hydrophobic interactions in a coiled-coil motif, as shown in Figure 8.

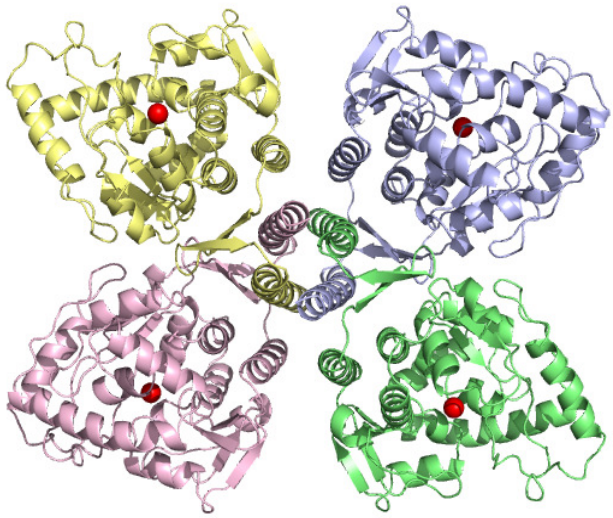


Figure 8: Structure of the catalytic and tetramerization domains of TH (PDB ID: 1TOH)(*63*)**. The coiled-coil motif connecting the monomers is seen in the center of the structure. Each monomer is a different color, and iron is shown as red spheres. The figure was made using PyMOL**(*50*)**.**

2.4 Mechanism

It is assumed that the AAAHs share a common catalytic mechanism(*14*, *26*). Here, the mechanism has been divided into three parts: The pre-reaction events including the effect of substrate and cofactor binding on the active site iron coordination, the oxygen activation step, which yields the hydroxylating intermediate, and the hydroxylation of the aromatic amino acid.

2.4.1 Substrate and cofactor binding

The effect of amino acid substrate and cofactor binding have been investigated using both crystal structures(*65*, *66*) and spectroscopic techniques(*67–70*) in anaerobic conditions to maintain iron in catalytically relevant +2 oxidation state. In the resting state, the active site iron in the AAAHs is 6 coordinated (6C) through 2 histidines, 1 glutamate and 3 water molecules (see Figure 9). Upon binding of either amino acid substrate or cofactor, no large structural changes are seen in overall structure or iron coordination. However, when both amino acid substrate and cofactor bind to complete the ternary complex, the enzyme structure closes in what appears to be a substrate triggered induced fit. The binding pocket of the substrate closes, leaving only a small entrance for oxygen to enter the active site. Furthermore, the iron coordination changes from 6C to 5 coordinated (5C) through a bidentate binding to glutamate and the loss of 2 water ligands. This yields an open coordination site for O_2 to react with iron to initiate the reaction.

By comparing PAH structures with cofactor (PDB ID: 1J8U), and cofactor and substrate analogue bound (PDB ID: 1MMK), the effect of substrate binding on the active site and overall structure can be investigated, as seen in Figure 9. The overall conformation changes from open to closed, along with a change in the active site iron coordination from 6C to 5C.

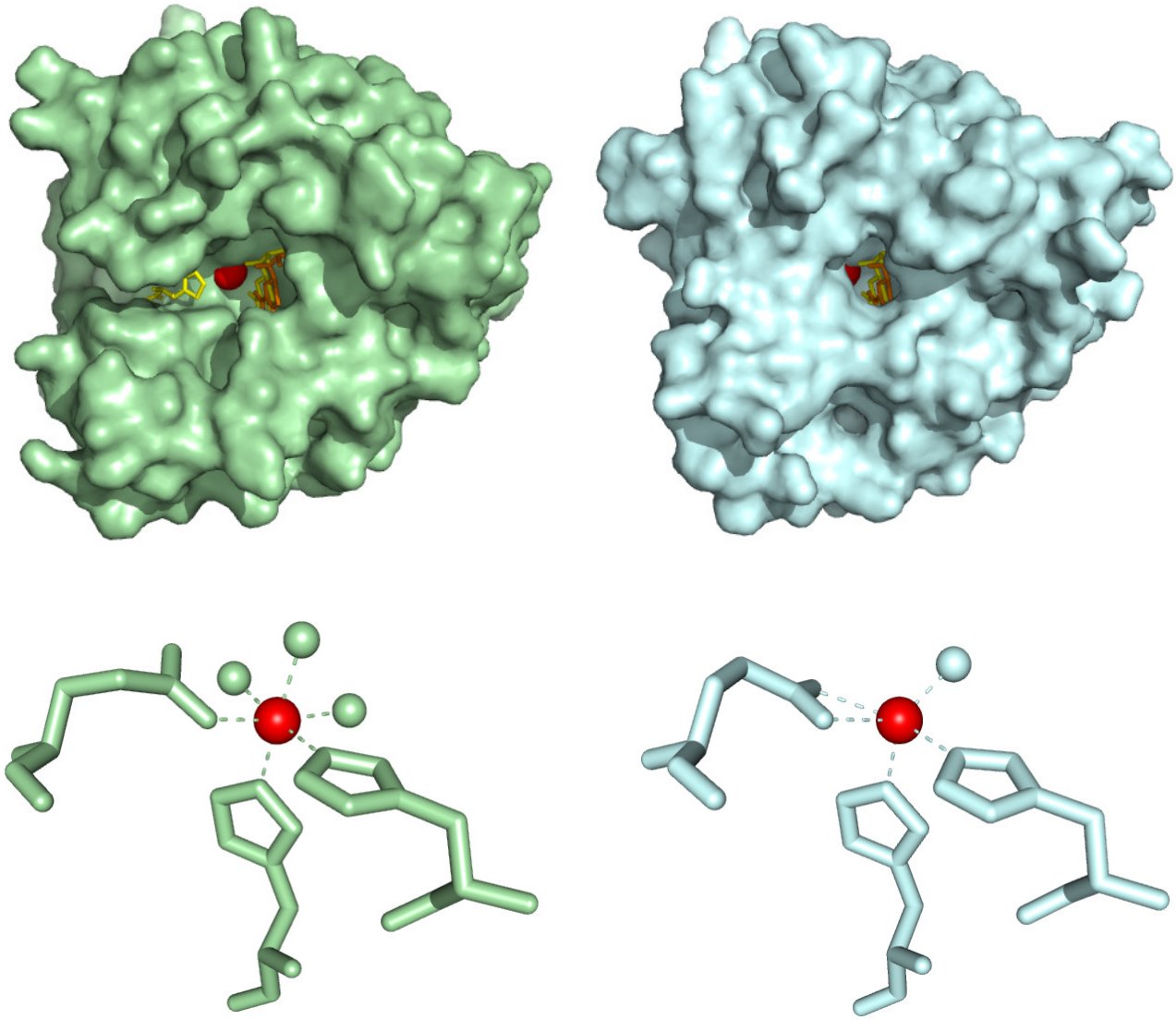


Figure 9: Comparison of the open (left) and closed (right) structures of PAH before and after substrate analogue binding. Left: Surface representation of PAH with Fe(II) (red sphere) and BH_4 bound (orange sticks) (PDB ID: 1J8U). Thienylalanine (substrate analogue) and BH_4 are aligned into the structure from the 1MMK structure by PyMOL and shown in yellow sticks. Below is the active site iron coordination. Right: Surface representation of PAH with Fe(II) (red sphere), BH_4 and thienylalanine bound (yellow sticks) (PDB ID 1MMK). BH_4 is aligned into the structure from the 1J8U structure by PyMOL and shown in orange sticks. Below is the active site iron coordination. The figure was made using PyMOL(*50*)**.**

Comparison of the two structures furthermore shows, that BH_4 is "pushed" approx. 2 Å closer to iron by movement of the backbone when substrate binds (see Figure 10). The movement of BH_4 pushes 2 waters away from the active site; HOH1251 which was coordinated to iron and Glu286, and HOH1142 which was coordinated to Glu286 and BH_4 before substrate binding. Glu286 now instead coordinates to BH_4 at two positions for a tighter binding.

HOH1253 also coordinated to iron is not directly pushed away by the BH_4 movement (distance between BH_4 and HOH1253 is 2.1 Å). Instead it might be removed with Tyr325 that swings out and could take the water with it. The remaining iron-coordinated water 1250 is the water closest to the position of the one iron-coordinated water in the ternary structure.

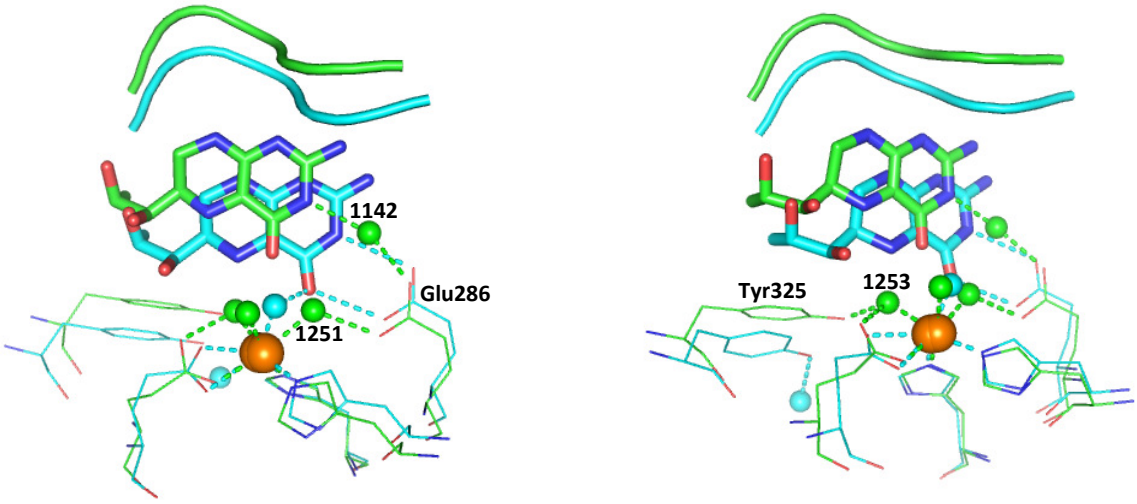


Figure 10: The movement of BH_4 upon substrate analogue binding in PAH. The structure with BH_4 bound (PDB ID: 1J8U) is shown in green and the structure with both BH_4 and substrate analogue bound (PDB ID: 1MMK) is shown in cyan. Fe(II) is shown as an orange sphere, and water as small spheres. Left: Interactions between Glu286 and water 1142 and 1251. Both waters are pushed away by the movement of BH_4. Right: Tyr325 swings out upon substrate binding and might take water 1253 with it, so glutamate can coordinate bidentate to Fe(II). The figure was made using PyMOL(*50*).

The catalytic mechanism and order of substrate binding have been studied for all three AAAHs by steady-state kinetics(*8*, *9*, *48*, *71–73*). It is concluded that the mechanism is sequential i.e. all three substrates must be bound in the active site before the reaction can occur. However, there is no consensus in the literature on the order of substrate binding or whether or not it is partially or fully ordered.

2.4.2 Oxygen activation and hydroxylation

The enzymatic reaction mechanism can be separated into two partial reactions. The first is the formation of the hydroxylating intermediate through activation of dioxygen and the second is the hydroxylation of the aromatic amino acid. The overall reaction mechanism shown in Figure 11 is general consensus in the literature(*14*), but specific reactions are still being discussed due to the lack of convincing experimental evidence.

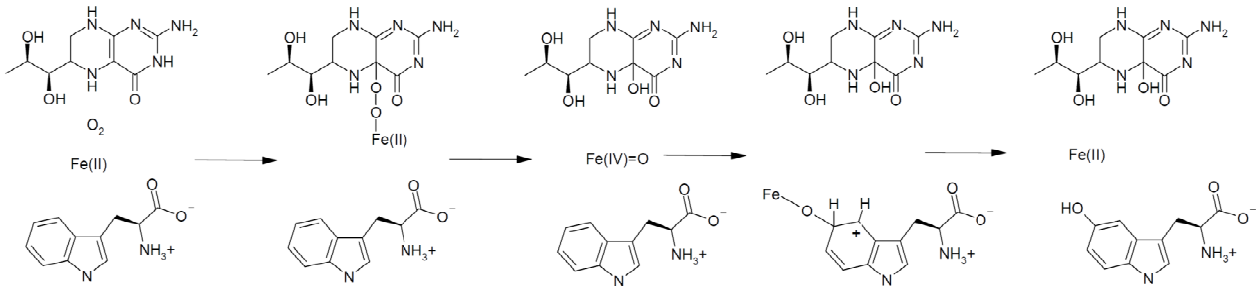


Figure 11: The proposed enzymatic reaction mechanism for TPH.

Theoretical investigations show that a ferryl oxy species (Fe(IV)=O) is a likely candidate for the hydroxylating intermediate(*74–78*), and recently direct spectroscopic evidence for a high-spin Fe(IV) intermediate, presumably Fe(IV)=O, was found for TH(*79*) and PAH(*80*). Fe(IV)=O is capable of arene hydroxylation(*81*) and is an intermediate in several mononuclear iron enzymes with facial triad

coordination(*75*, *82*). The Fe(IV)=O intermediate is formed from heterolytic cleavage of a Fe(II) peroxypterin bridge. However, disagreements exist on whether oxygen reacts with iron or pterin before formation of the peroxy bridge(*74*, *83*, *84*).

After formation of the hydroxylating intermediate, the first step in the hydroxylation of the amino acid is believed to be the electrophilic attack of Fe(IV)=O on the aromatic C-atom as shown in Figure 11. The product is an arenium cation intermediate. This step has been supported by experimental evidence(*26*, *85*–*87*). The formation of the arenium cation is followed by a NIH-shift, which has been experimentally observed for all AAAHs(*85*, *88*–*90*). The hydroxylation reaction mechanism proposed is furthermore supported by theoretical calculations(*77*, *91*, *92*).

The hydroxylation of tryptophan is the rate-limiting step in the enzymatic turnover by TPH(*26*, *86*, *91*). On the contrary, the oxygen activation step is believed to be the rate-limiting step of TH and PAH(*71*, *83*, *91*, *93*).

Even though several steps in the reaction mechanism have been studied, there are still many unknown factors which need to be investigated in order to present a clear picture of the mechanism.

2.5 Nomenclature

In this thesis, the different TPH variants have been given abbreviations, which follow the general rules described here. If the variant is fused to an affinity tag, the abbreviation for the affinity tag is written first followed by "-". Affinity tags featured in this project includes gluthathione S-transferase (GST) and maltose binding protein (MBP). The first letter (or first two letters) refers to the domains. The regulatory domain is assigned "r", the catalytic domain "c", and the tetramerization domain "t". For the full-length enzymes no letter is included. The next letter in italic refers to the species that the protein sequence originates from. "*h*" refers to *homo sapiens* (human) and "*g*" to *gallus gallus* (chicken). The protein name is abbreviated in capital letters, such as "TPH1". Examples of the most commonly used abbreviations are given in Table 1.

Table 1: Overview of the abbreviations used in this thesis for TPH variants.

Abbreviation	Affinity-tag	Domains	Species	Protein
c*g*TPH1	-	Catalytic domain	Chicken	TPH1
GST-c*h*TPH1	GST	Catalytic domain	Human	TPH1
MBP-*h*TPH1	MBP	Full-length	Human	TPH1
MBP-rc*h*TPH1	MBP	Regulatory and catalytic domains	Human	TPH1
c*h*TPH2	-	Catalytic domain	Human	TPH2
ct*h*TPH2	-	Catalytic and tetramerization domains	Human	TPH2
***h*TPH2**	-	Full-length	Human	TPH2
MBP-*h*TPH2	MBP	Full-length	Human	TPH2
MBP-rc*h*TPH2	MBP	Regulatory and catalytic domains	Human	TPH2

3 Project background and strategy

The Metalloprotein Chemistry & Engineering group at DTU Chemistry started working with TPH in 2003. Since then, a vast amount of experimental data has been accumulated. All these data along with the contributions from other groups studying TPH gives the knowledge platform upon which this project is build. In this chapter, a short review of the project background for the three main topics (purification, crystallization and mechanistic studies) will be given, and the strategies followed in this project will be presented.

From the very beginning, purification of full-length TPH proved difficult. The soluble amount of protein after expression in *Escherichia coli* (*E. coli*) was very low, and the protein easily precipitated and would not bind to the columns tested(*9*). Thus, a new strategy was developed were each of the three domains were studied, first individually and later in different combinations, to achieve knowledge on which parameters were important for stability. The purification of the catalytic domain was rather simple, as this variant was expressed in large amounts and was very soluble. The method of purification included a capture step using anion-exchange chromatography and a polishing step using size-exclusion chromatography(*13*). The yields were high and the product rather stable. c*h*TPH2 was found to be more stable than c*h*TPH1, and could be concentrated to a much higher concentration (2.5 mM vs. 0.2 mM). The inclusion of the tetrameric domain did not complicate the purification process and ct*h*TPH2 could be purified using the same method as for the catalytic domain(*11*). ct*h*TPH1 could be expressed however in a relatively low yield, but was found soluble(*94*). A purification method for ct*h*TPH1 has not been developed. Experiments with variants containing the regulatory domain gave either insoluble protein, or partly soluble protein that existed in large aggregates and did not bind to ion-exchange columns or eluted in several different peaks(*9*, *94*). No suitable method for purification of any variant containing the regulatory domain was developed until recently, when a successful purification method for GST-r*h*TPH1 was found(*10*). The method includes a high concentration of salt in the buffers used, in order to stabilize the regulatory domain and prevent oligomerization and/or aggregation. This important result is the base for developing purification strategies for the full-length variants of TPH1 and TPH2. The strategy when developing new purification methods is that they should be as simple as possible, meaning the fewer steps the better. If the yield is low, any critical steps should be identified and improvements or replacement of the critical step should be tested. If the same purification method can be used for several different variants, this is preferred, as it limits the number of protocols and columns in use. Furthermore, the product from the purification should be tested for purity (using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and mass spectrometry) and activity. The stability of the product can also be tested.

As mentioned in the introduction, only very limited information on the structure of TPH is available. The only crystal structures available are of the catalytic domains of isoform 1. Crystal structures of both c*g*TPH1

and c*h*TPH1 have been solved in the Metalloprotein Chemistry & Engineering group(*10*, *12*). Several attempts have been made to crystallize c*h*TPH2 and ct*h*TPH2, but all were unsuccessful(*9*, *11*). In my master thesis a low-resolution structure of ct*h*TPH2 was determined using small angle X-ray scattering, and this showed a structure very similar to the crystal structure of tetrameric TH(*11*). The purification products of c*h*TPH2 and ct*h*TPH2 were also tested using different scattering techniques, and it was found that c*h*TPH2 was polydisperse, while ct*h*TPH2 was monodisperse, also after freezing and thawing(*11*). Further explanations for the lack of crystallization results could be the high content of glycerol in the purification buffers (10 (w/v) %), which might have a negative effect on the crystallization process. Crystallization is indeed a hit-or-miss science, and great effort does far from always lead to great results. There is a vast amount of parameters which can be varied in crystallization, both for the purification product (purification method, buffer, concentration, freezing/thawing etc), and for the crystallization conditions (precipitant, temperature, additives, substrates, drop size, drop type (hanging, sitting etc)). Generally, if no hits are found when testing two different screens at two different temperatures and two different protein concentrations, then the protein sample should be checked for monodispersity, purity etc(*95*). The goal in this project regarding crystallization was to find crystal structures for any variants of TPH that have not been crystallized before, or of known structures in new combinations with substrates. The strategy involved getting pure protein samples for as many variants as possible, so screens could be set up. For crystallization, both random variation (sparse matrix screens) as well as variations of specific parameters (fine-tuning), were tested.

Since the achievement of pure and stable products of TPH in the group is relatively new, the characterization results are also limited. However, c*h*TPH1, c*h*TPH2 and c*g*TPH1 have been kinetically characterized by determining the apparent Michaelis menten parameters, K_M and V_{max}, for all 3 substrates(*13*, *96*). c*h*TPH1 and c*g*TPH1 were inhibited by tryptophan at high concentrations, while c*h*TPH2 did not show such substrate inhibition. Moreover, the K_M value for the cofactor BH_4 differed significantly between the isoforms, as $K_{M,BH4}$ was 10-fold higher for c*h*TPH1 than for c*h*TPH2. These isoform specific differences are rather surprising, as the catalytic domains have a high sequence identity. Whether these differences in K_M arise from difference in association constants or different reaction rates is unknown. So far, the reaction mechanism of TPH has not been studied in the Metalloprotein Chemistry & Engineering group. The strategy in this project regarding mechanistic studies was to gain as much information as possible on the structure-functional relationship and reaction mechanism of TPH using several different techniques available.

4 Purification

In this chapter the purification methods developed for a wide range of TPH variants will be presented. The characterization of the final products will also be shown and discussed. For most variants, only the final purification procedure is presented, even though several purifications have been performed during development and optimization of the method.

First, an introduction to previous work with purification of TPH is given along with a strategy for the work performed in this project. Then the experimental procedures are presented, followed by results and discussion. The stability of some of TPH variants was studied by differential scanning fluorimetry (DSF), and these results will also be presented. Finally, a comparison between the yields, solubility, and activities of the different variants will be made.

4.1 Introduction and background

A requirement for thorough characterization is a pure and stable protein sample. This is usually achieved by expression in recombinant systems followed by purification. Alternatively, the protein can be purified from the original source.

For many years work have been performed on purification of full-length variants of TPH from recombinant systems. Both PAH and TH have been purified, but the inherent instability of TPH and its tendency to aggregate makes it very hard to work with. The Metalloprotein Chemistry & Engineering group has succeeded in developing purification procedures for several truncated variants containing the catalytic and tetramerization domains, as described in the previous chapter. These include purification of c*h*TPH1(*13*), c*h*TPH2(*13*), c*g*TPH1(*96*), and ct*h*TPH2(*9*). The products have been characterized kinetically(*9*, *13*), and a crystal structure of c*g*TPH1 was also solved(*12*). Recently, a purification method for the regulatory domain (r*h*TPH1) was developed using GST as affinity tag and high salt concentrations in the buffer systems(*10*).

From the literature it is known that other research groups have purified full-length human TPH fused to affinity tags(*51*, *97*–*101*). Especially the group of Jan Haavik has in recent years developed promising purification procedures for both human TPH1 and TPH2(*51*, *98*, *102*). In all studies, TPH becomes unstable and precipitate or aggregate after cleavage from the affinity tags resulting in no or very low yields. For TPH2, the protein solution is not pure after cleavage(*51*). Only one successful purification for full-length human TPH1 has been reported, using a pterin-agarose column(*103*). Some discrepancies are however seen in the results from this purification method. One example is a 10-fold increase in activity upon purification, which is explained by the presence of unknown endogenous inhibitors of TPH in the crude extract(*103*). Furthermore, when the method with pterin-agarose columns was tested in our lab it did not succeed in binding TPH(*104*). Thus, the evidence of a pure and stable full-length TPH product remains elusive.

The objective for the purification part of this project was to develop purification methods for full-length *h*TPH1 and *h*TPH2. The initially developed procedures for the catalytic domains of TPH consist of an anion-exchange step followed by a size-exclusion step. However, as earlier studies have shown that the regulatory domain of *h*TPH1 requires high salt concentration for stabilization(*10*), new methods should be developed

that allows for this requirement. Thus, ion-exchange is excluded. Affinity purification was chosen as the preferred method due to high selectivity, fast purification and the possibility for high salt concentration. Affinity purification was combined with size-exclusion chromatography in order to allow for aggregates to be removed (*vide infra*, see section 4.3.2.2) and for effective removal of the eluent before the second affinity step.

For monomer TPH variants gluthathione S-transferase (GST) can be used as affinity tag. However, due to the common dimerization of GST(*105*) it is not suitable as affinity tag for tetrameric variants of TPH. Thus, maltose binding protein (MBP) was chosen as the preferred affinity tag for all TPH variants except c*h*TPH1 which was fused to GST. For cleavage of affinity tag from target protein, human rhinovirus 3C protease (3CP) was chosen as this allows for quick and specific cleavage(*106*). 3CP recognizes the amino acid sequence Leu-Glu-Val-Leu-Phe-Gln↓Gly-Pro, cleaving between the Gln and Gly residues(*107*). Thus, two extra amino acids will be left on the target protein after cleavage. 3CP displays maximum activity at 4°C, hence the cleavage can be performed at low temperatures ensuring protein stability. Furthermore, the protease can be fused to affinity tags such as GST or MBP and removed from the solution by affinity purification in an easy procedure.

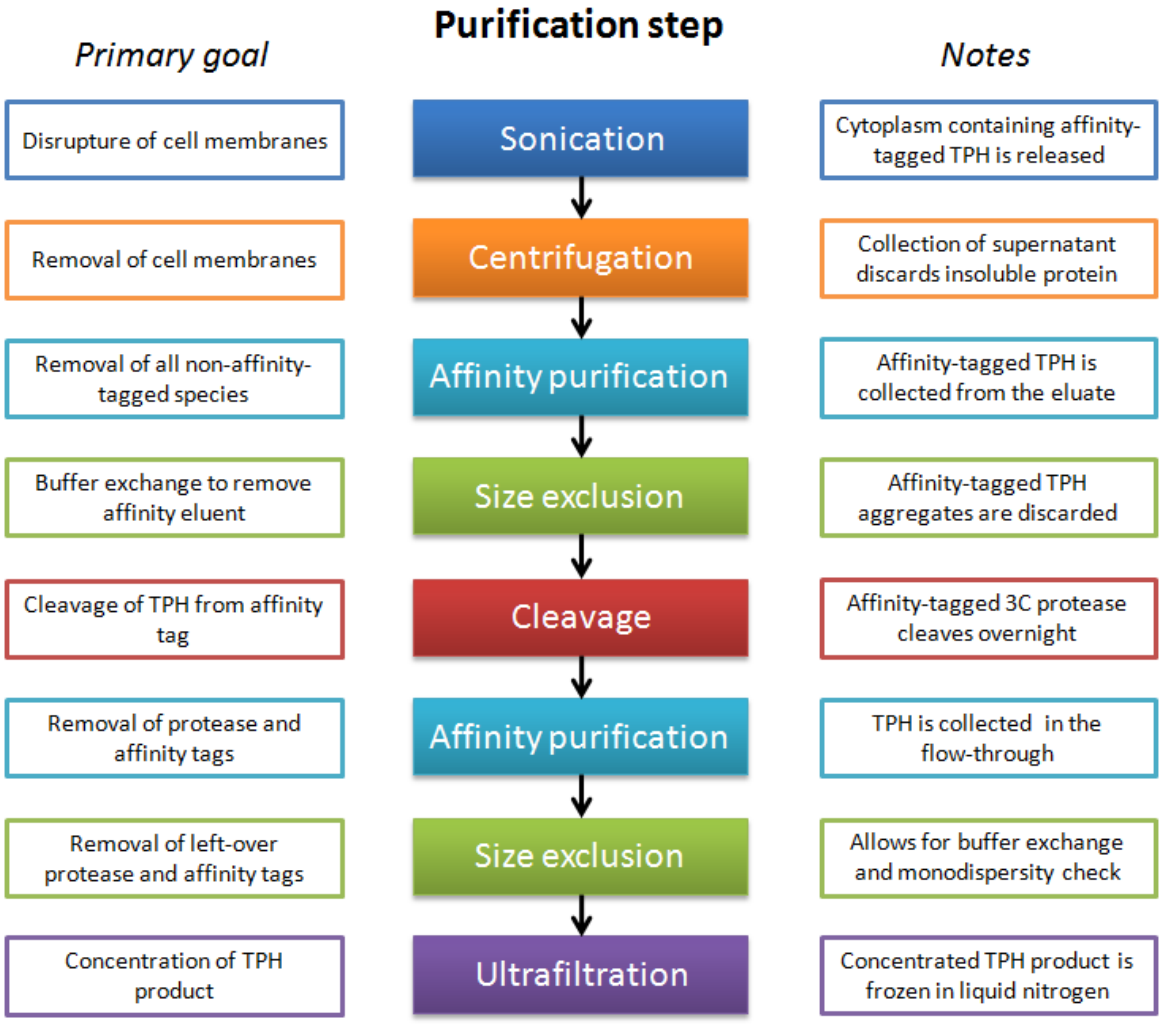


Figure 12: Flow chart of the general purification method for affinity-tagged TPH variants. In the middle each purification step is shown, to the left the primary goal of the purification step is shown, and to the right any notes on the purification step are shown.

The genes for affinity-tagged TPH variants were constructed encoding an affinity-tag at the N-terminal followed by a linker (MBP variants only), the protease recognition sequence, and TPH. For the affinity-tagged 3CP no recognition sequence was included. For complete amino acid sequences, see appendix 9.1.
A flow chart of the overall purification method for TPH variants using affinity tags can be seen in Figure 12. The method includes sonication followed by the first affinity and size-exclusion steps. Then the affinity-tagged TPH variant is cleaved by 3CP overnight, and the product is purified by the second affinity and size-exclusion steps. Finally the product is concentrated using ultrafiltration.

4.2 Experimental procedures

In this section the experimental procedures for expression and purification of different TPH variants and proteases will be presented. The methods are described for 3 tubes of cells (corresponding to 1.95 L culture) whenever possible, but can easily be downscaled to one tube by using smaller columns.
The construction of genes was performed by other members of the group, and will not be described. The expression and purification methods for c*h*TPH2 and ct*h*TPH2 were developed by Nielsen(*9*) and myself(*11*) in earlier projects, and have only been slightly modified in this project. The purification method for *h*TPH2 without affinity tag was developed in my master thesis (*11*), and was also only slightly modified in this project. The modifications include purification without glycerol in the buffers and without reducing conditions (addition of dithiothreitol (DTT) and argon flow).
During my PhD I have supervised several bachelor-students, who have contributed to the experimental work. Barbara Gonzalez, Stine Jørgensen and Mads Friis worked on purification, stability and kinetic analysis of *h*TPH1, Maria Oettinger on purification and mass spectrometry of c*h*TPH1, Katrine Bünger on mass spectrometry of c*h*TPH2 and Gro Nyborg on purification of *h*TPH2 using MBP as affinity tag.
All chemicals used are of analytical grade and the term “water” refers to 18.2 MΩ/cm water from a Milli-Q synthesis A10 Q-Gard system (Millipore). During purification, all solutions containing TPH were kept on ice or at 4 °C except during the chromatographic steps, which were performed at room temperature. High-performance liquid chromatography (HPLC) was performed using columns and an ÄKTApurifier 100 from GE Healthcare. Protein concentration was measured by ultraviolet-visible spectroscopy (UV-Vis) using a Nanodrop 1000 spectrophotometer from Thermo Scientific. Theoretical extinction coefficients at 280 nm (ε_{280nm}) were determined according to reference (*108*) using the ProtParam program from the Expasy webpage(*109*).

4.2.1 Expression

E. coli BL21(DE3) strains expressing the target-protein from a pET26 expression vector were taken from frozen glycerol stocks and streaked into single colonies on a Luria Bertani (LB) agar plate containing either 30 µg/mL kanamycin sulfate (30kan) or 100 µg/mL ampenicillin (100amp), according to Table 2. The plate was incubated overnight at 37 °C. A single colony was inoculated in 50 mL media in a 250 ml shake flask. This pre-culture was incubated at 37 °C and 250 rpm for approx. 4 hours until an optical density at 600 nm (OD_{600}) of 0.6-1.0 was reached. The cells were transferred to a 50 mL polypropylene tube and sedimented by centrifugation at 4 °C and 1800 x g for 10 min. The cells were resuspended in 50 mL fresh media. 6.5 mL cell resuspension was used to inoculate 650 mL media in a 2 L triple baffled shake flask. The culture was incubated at 30 °C and 250 rpm. For most target-proteins, the temperature was lowered to 20 °C after approx. 3½, but for some it was kept at 30 °C, see Table 2. When an OD_{600} of 0.4-0.6 was attained, the target-protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final

concentration of 0.1 mM. For some target-proteins a freshly made sterile solution of $(NH_4)_2Fe(II)(SO_4)_2$ was added to a final concentration of 0.2 mM. The cells were incubated overnight at 20 °C/30 °C and 250 rpm. The cells were harvested by centrifugation at 4 °C and 3000 x g for 15 min. The cells from 650 ml culture were resuspendend in 25 ml of the appropriate buffer (see Table 2), transferred to 50 ml polypropylene tubes and centrifuged again at 4 °C and 3000 x g for 15 min. The supernatant was discarded and the cells were stored at -80 °C until further use.
The expression of *h*TPH2 using auto-inducing media is described in reference (*11*).

Table 2: Expression details of strains used in this project. HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.

Variant	Internal ID number	Media	Expression temperature	Addition of Fe(II)	Harvesting buffer
GST-3CP	HC2510	LB/100amp	30 °C	No	1xPBS, pH 7.3
MBP-3CP	HC2647	LB/30kan	30 °C	No	20 mM Tris/H_2SO_4, 70 mM $(NH_4)_2SO_4$, pH 8.0
GST-c*h*TPH1	HC2641	LB/100amp	20 °C	Yes	1xPBS/NaOH, 200 mM $(NH_4)_2SO_4$, pH 8.5
MBP-*h*TPH1	HC2595	LB/30kan	20 °C	Yes	20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0
MBP-rc*h*TPH1	HC2750	LB/30kan	20 °C	Yes	20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0
c*h*TPH2	HC1168	LB/30Kan	20 °C	No	20 mM Bis Tris Propane/HCl, pH 7.2
ct*h*TPH2	HC2334	LB/30Kan	20 °C	No	20 mM Bis Tris Propane/HCl, pH 7.2
MBP-*h*TPH2	HC2878	LB/30kan	20 °C	Yes	20 mM HEPES/NaOH, 300 mM $(NH_4)_2SO_4$, pH 7.0
MBP-rc*h*TPH2	HC2875	LB/30kan	20 °C	Yes	20 mM HEPES/NaOH, 300 mM $(NH_4)_2SO_4$, pH 7.0

4.2.2 Purification of proteases and cleavage test

A simple purification of the affinity-tagged proteases was performed as described below. The procedure involved only one step, as the purity of the product is not of great importance.

4.2.2.1 Purification of GST-3CP

Cells expressing GST-3CP (HC2510) from 3 x 650 ml culture were resuspended in 1xPBS, pH 7.3 to a total volume of approx. 3 x 40 mL. The cells were lysed on ice by sonication for 3 x 30 sec using a Satorius Labsonic P at 80 % amplitude. The cell extract was centrifuged in high-speed centrifuge tubes at 18000 x g and 4 °C for 25 min. The supernatant was collected and filtrated through a 0.45 µm GHP Acrodisc GF syringe filter. The filtrated sample was loaded (1 ml/min) onto a 20 mL GSTPrep FF 16/10 column, which was equilibrated with 1xPBS, pH 7.3. The protein was eluted with 50 mM Tris/H_2SO_4, 50 mM $(NH_4)_2SO_4$, 10 mM reduced glutathione, pH 8.0. Reduced glutathione was added just before use, followed by pH readjustment with H_2SO_4. The eluted protein was collected and exchanged into a storage buffer containing 50 mM Tris/H_2SO_4, 50 mM $(NH_4)_2SO_4$, 1 mM DTT, 10 mM ethylenediaminetetraacetic acid (EDTA), 20 % glycerol, pH 8.0. DTT was added just before use. The buffer exchange was performed by concentrating the protein using

ultrafiltration on an Amicon stirred pressure cell with a PL-3 membrane and subsequently diluting it with storage buffer. This was repeated three times. The concentration of GST-3CP was determined using ε_{280nm} = 64290 M^{-1} cm^{-1}. The protein was stored at -20 °C.

4.2.2.1.1 Cleavage test

A cleavage test was performed in order to investigate the optimal ratio between GST-c*h*TPH1 and GST-3CP for complete cleavage. GST-c*h*TPH1 was purified as described in section 4.2.3.1 and mixed in different molar ratios with GST-3CP purified as described in section 4.2.2.1. The solution was left overnight at 4 °C The molar ratios tested were 1:1, 1:2, 1:4, 1:7, 1:15, 1:29, and 1:59 (GST-3CP:GST-c*h*TPH1). To allow for a direct comparison on SDS-PAGE, all samples had the same total volume, and the concentration of GST-c*h*TPH1 was also kept constant.

4.2.2.2 Purification of MBP-3CP

Cells expressing MBP-3CP (HC2647) from 3 x 650 ml culture were resuspended in 20 mM Tris/H_2SO_4, 70 mM $(NH_4)_2SO_4$, pH 8.0 to a total volume of approx. 3 x 40 mL. The cells were lysed on ice by sonication for 3 x 30 sec. The cell extract was centrifuged in high-speed centrifuge tubes at 18000 x g and 4 °C for 25 min. The supernatant was collected and filtrated. The filtrated sample was loaded (2.5 ml/min) onto a 20 mL Dextrin Sepharose HP 16/10 MBP column, which was equilibrated with 20 mM Tris/H_2SO_4, 70 mM $(NH_4)_2SO_4$, pH 8.0. The protein was eluted with 20 mM Tris/H_2SO_4, 70 mM $(NH_4)_2SO_4$, 10 mM maltose, pH 8.0. The eluted protein was collected and exchanged into a storage buffer containing 50 mM Tris/H_2SO_4, 50 mM $(NH_4)_2SO_4$, 1 mM DTT, 10 mM EDTA, 20 % glycerol, pH 8.0. DTT was added just before use. The buffer exchange was performed by concentrating the protein using ultrafiltration and subsequently diluting it with storage buffer. This was repeated three times. The concentration of MBP-3CP was determined using ε_{280nm} = 87780 M^{-1} cm^{-1}. The protein was stored at -20 °C.

4.2.2.2.1 Cleavage test

A cleavage test was performed in order to investigate the optimal ratio between MBP-*h*TPH1 and MBP-3CP for complete cleavage. MBP-*h*TPH1 was purified as described in section 4.2.3.2 and mixed in different molar ratios with MBP-3CP purified as described in section 4.2.2.2. The solution was left overnight at 4 °C. The molar ratios tested were 1:3, 1:6, 1:50, 1:100, and 1:200 (MBP-3CP:MBP-*h*TPH1). To allow for a direct comparison on SDS-PAGE, all samples had the same total volume, and the concentration of MBP-*h*TPH1 was also kept constant.

4.2.3 Purification of TPH1

Development of the purification procedure started with purification of GST-c*h*TPH1. The original purification procedure for c*h*TPH1 containing an anion-exchange step followed by a gel-filtration step was found insufficient, as mass spectrometry revealed impurities in the final product(*10*). Thus, a new method involving affinity tag was developed for c*h*TPH1. Later, the method was extended to *h*TPH1, where MBP was used as affinity tag. The experimental procedures for the final purification procedures will be described in this section.

*4.2.3.1 Purification of GST-tagged c*h*TPH1*

GST-c*h*TPH1 was purified according to the general procedure shown in Figure 12 including an affinity step followed by gel-filtration, cleavage, a second affinity step and a second gel-filtration. The first gel-filtration was included to remove reduced glutathione from the solution, which would otherwise disturb the second

affinity step. The last gel-filtration step was included as a polishing step to remove any left-over GST or GST-3CP and to allow for buffer exchange and monodispersity check.

Cells expressing GST-c*h*TPH1 (HC2641) from 3 x 650 ml culture were resuspended in 1xPBS/NaOH, 200 mM $(NH_4)_2SO_4$, pH 8.5 to a total volume of approx. 3 x 40 mL. The cells were lysed on ice by sonication for 3 x 30 sec. The cell extract was centrifuged in high-speed centrifuge tubes at 18000 x g and 4 °C for 35 min. The supernatant was collected and filtrated. The filtrated sample was loaded (2.0 ml/min) onto a 20 mL GSTPrep FF 16/10 column, which was equilibrated with 1xPBS/NaOH, 200 mM $(NH_4)_2SO_4$, pH 8.5. The protein was eluted with 100 mM Tris/NaOH, 200 mM $(NH_4)_2SO_4$, 10 mM reduced glutathione, pH 8.5. Reduced glutathione was added just before use, followed by pH readjustment. The eluted protein was collected and concentrated by ultrafiltration to approx. 5 mL. The concentrated sample was filtrated and loaded onto a HiLoad Superdex 200 26/60 prep grade column, which was equilibrated with 1xPBS/NaOH, 200 mM $(NH_4)_2SO_4$, pH 8.5. The fractions containing GST-c*h*TPH1 were collected and the concentration was determined using ε_{280nm} = 76210 M^{-1} cm^{-1}.The cleavage was performed by mixing GST-3CP with GST-c*h*TPH1 in a molar ratio of 1:7. The solution was left overnight at 4 °C. The cleaved sample was filtrated and loaded (1 ml/min) onto a 20 mL GSTPrep FF 16/10 column, which was equilibrated with 1xPBS/NaOH, 200 mM $(NH_4)_2SO_4$, pH 8.5. c*h*TPH1 was collected from the flow-through, and the concentration was determined using ε_{280nm} = 33350 M^{-1} cm^{-1}. Fe(II) was added in 1.5 x molar excess from a freshly made 100 mM $(NH_4)_2Fe(II)(SO_4)_2$ solution, and the sample was incubated for 15 min. The sample was concentrated by ultrafiltration to approx. 5 mL. The concentrated sample was filtrated and loaded onto a HiLoad Superdex 75 26/60 prep grade column, which was equilibrated with 1xPBS/NaOH, 200 mM $(NH_4)_2SO_4$, pH 8.5. The fractions containing c*h*TPH1 were collected and concentrated using ultrafiltration. The final product was frozen in liquid nitrogen and stored at -80 °C.

4.2.3.2 Purification of MBP-tagged h*TPH1*

MBP-*h*TPH1 was purified according to the general procedure shown in Figure 12.

Cells expressing MBP-*h*TPH1 (HC2595) from 3 x 650 ml culture were resuspended in 20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0 to a total volume of approx. 3 x 40 mL. The cells were lysed on ice by sonication for 3 x 30 sec. The cell extract was centrifuged in high-speed centrifuge tubes at 18000 x g and 4 °C for 35 min. The supernatant was collected and filtrated. The filtrated sample was loaded (2.5 ml/min) onto a 20 mL Dextrin Sepharose HP 16/10 MBP column, which was equilibrated with 20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0. The protein was eluted with 20 mM Tris/ H_2SO_4, 300 mM $(NH_4)_2SO_4$, 10 mM maltose, pH 8.0. The eluted protein was collected and concentrated by ultrafiltration to approx. 5 mL. The concentrated sample was filtrated and loaded onto a HiLoad Superdex 200 26/60 prep grade column, which was equilibrated with 20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0. The fractions containing MBP-*h*TPH1 were collected and the concentration was determined using ε_{280nm} = 99700 M^{-1} cm^{-1}.The cleavage was performed by mixing MBP-3CP with MBP-*h*TPH1 in a molar ratio of 1:7. The solution was left overnight at 4 °C. The cleaved sample was filtrated and loaded (1.5 ml/min) onto a 20 mL Dextrin Sepharose HP 16/10 MBP column, which was equilibrated with 20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0. *h*TPH1 was collected from the flow-through, and the concentration was determined using ε_{280nm} = 33350 M^{-1} cm^{-1}. Fe(II) was added in 1.5 x molar excess from a freshly made 100 mM $(NH_4)_2Fe(II)(SO_4)_2$ solution, and the sample was incubated for 15 min. The sample was concentrated by ultrafiltration to approx. 5 mL. The concentrated sample was filtrated and loaded onto a HiLoad Superdex 200 26/60 prep grade column, which was equilibrated with 20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0. The fractions containing *h*TPH1 were

collected and concentrated by ultrafiltration. The final product was frozen in liquid nitrogen and stored at -80 °C.
The purification was also performed with 50 mM $(NH_4)_2SO_4$, 300 mM Na_2SO_4 and 900 mM $(NH_4)_2SO_4$ instead of 300 mM $(NH_4)_2SO_4$ in all buffers.
The solubility of GST-*h*TPH1 was tested in 20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0 using the first part of the procedure for MBP-*h*TPH1 described above.

*4.2.3.3 Purification of MBP-tagged rc*h*TPH1*

MBP-rc*h*TPH1 was purified using the exact same procedures as for MBP-*h*TPH1 described in section 4.2.3.2. The concentration was determined using the same extinction coefficients as for MBP-*h*TPH1, as the tetramerization domain in MBP-*h*TPH1 does not contribute to the absorption at 280 nm due to the lack of tryptophan, tyrosine and cystine residues.

4.2.4 Purification of TPH2

The purification procedure for c*h*TPH2, ct*h*TPH2, and *h*TPH2 without affinity tag are quite different from the previous described methods. They consist only of an anion-exchange step and a gel-filtration step, and can be performed in one day. The experimental procedures for purification of both non affinity-tagged and affinity-tagged TPH2 variants will be described here.

*4.2.4.1 Purification of c*h*TPH2*

Cells expressing c*h*TPH2 (HC1168) from 3 x 650 ml culture were resuspended in 20 mM Bis Tris Propane/HCl, pH 7.2 to a total volume of approx. 3 x 40 mL. The cells were lysed on ice by sonication for 3 x 30 sec. The cell extract was centrifuged in high-speed centrifuge tubes at 18000 x g and 4 °C for 25 min. The supernatant was collected and filtrated. The filtrated supernatant was diluted with ice cold water to a conductivity of ~2 mS/cm. The diluted supernatant was loaded (7ml/min) onto a Q Sepharose High Performance 26/10 column, which was equilibrated in 20 mM Bis Tris Propane/HCl, pH 7.2. For elution, a linear gradient of 0-4 % 20 mM Bis Tris Propane, 0.8 M $(NH_4)_2SO_4$, pH 7.2 (0-32 mM $(NH_4)_2SO_4$) over one column volume (CV) was applied to the column, followed by a gradient from 4-12 % 20 mM Bis Tris Propane, 0.8 M $(NH_4)_2SO_4$, pH 7.2 (32-96 mM $(NH_4)_2SO_4$) over 5 CV. The fractions containing c*h*TPH2 were collected and the concentration was determined using ε_{280nm} = 39310 M^{-1} cm^{-1}. Fe(II) was added in 1.5 x molar excess from a freshly made 100 mM $(NH_4)_2Fe(II)(SO_4)_2$ solution, and the sample was incubated for 15 min. The sample was concentrated by ultrafiltration to approx. 5 mL. The concentrated sample was filtrated and loaded onto a HiLoad Superdex 75 26/60 prep grade column, which was equilibrated with 20 mM HEPES/NaOH, 100 mM $(NH_4)_2SO_4$, pH 7.2. The fractions containing c*h*TPH2 were collected and concentrated using ultrafiltration. The final product was frozen in liquid nitrogen and stored at -80 °C.
The gel-filtration step was performed in several different buffers including 300 mM NH_4OAc and 20 mM HEPES/NaOH, 200 mM NaCl, pH 7.2. The purification has also been performed with 5 (w/v)% glycerol in all buffers.

*4.2.4.2 Purification of ct*h*TPH2*

cthTPH2 was purified using the same method as for c*h*TPH2 described in section 4.2.4.1, with the following exceptions: I) A linear gradient of 5-25 % 20 mM Bis Tris Propane, 0.8 M $(NH_4)_2SO_4$, pH 7.2 (0-200 mM $(NH_4)_2SO_4$) over 10 CV was applied to the anion-exchange column. II) ε_{280} = 40800 M^{-1} cm^{-1} for ct*h*TPH2. III) A HiLoad Superdex 200 26/60 prep grade column was used in the gel-filtration step.

The gel-filtration step was performed in several different buffers including 0.3 M NH_4OAc /NH_3, pH 7.2 and 10 mM HEPES/NaOH, 200 mM NaCl, pH 7.2 The purification has also been performed with 10 (w/v)% glycerol in all buffers for an extra stabilizing effect.

4.2.4.3 Purification of h*TPH2*

The purification procedure for *h*TPH2 was developed in reference (*11*), and generally follows the same procedure as for c*h*TPH2 and ct*h*TPH2. Only small modifications to the procedure have been performed in this project.

*h*TPH2 was purified using the same methods as for c*h*TPH2 described in section 4.2.4.1, with the following exceptions: I) A linear gradient of 0-35 % 20 mM Bis Tris Propane, 0.8 M $(NH_4)_2SO_4$, pH 7.2 (0-280 mM $(NH_4)_2SO_4$) over 10 CV was applied to the anion-exchange column. II) 2M $(NH_4)_2SO_4$ solution was added to the protein sample to a final concentration of 100 mM after the anion-exchange step. III) ε_{280} = 53290 M^{-1} cm^{-1} for *h*TPH2. IV) A HiLoad Superdex 200 26/60 prep grade column was used in the gel-filtration step.

*4.2.4.4 Purification of MBP-tagged rc*h*TPH2*

MBP-rc*h*TPH2 was purified using the same procedures as for MBP-*h*TPH1 described in section 4.2.3.2, with the following exceptions: I) Only one tube of cells was used for the purification. II) The buffers used were 20 mM HEPES/NaOH, 300 mM $(NH_4)_2SO_4$, pH 7.0 and 20 mM HEPES/NaOH, 300 mM $(NH_4)_2SO_4$, 10 mM maltose, pH 7.0. III) The MBP column used was a MBPTrap HP 5 mL column. IV) ε_{280} = 118150 M^{-1} cm^{-1} for MBP-rc*h*TPH2 and ε_{280} = 51800 M^{-1} cm^{-1} for rc*h*TPH2.

4.2.4.5 Purification of MBP-tagged h*TPH2*

MBP-*h*TPH2 was purified using the same procedures as for MBP-*h*TPH1 described in section 4.2.3.2, with the following exceptions: I) The buffers used were 20 mM HEPES/NaOH, 300 mM $(NH_4)_2SO_4$, pH 7.0 and 20 mM HEPES/NaOH, 300 mM $(NH_4)_2SO_4$, 10 mM maltose, pH 7.0. II) ε_{280} = 119640 M^{-1} cm^{-1} for MBP-*h*TPH2 and ε_{280} = 53290 M^{-1} cm^{-1} for *h*TPH2.

4.2.5 Characterization methods

The experimental procedures for the characterization methods used in this chapter are presented here.

4.2.5.1 SDS-PAGE

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was carried out as described here. Tris-HCl gels from Bio-Rad were used and depending on the target protein 7.5 %, 12 % or 18 % gels were chosen for optimal separation. The gels were prepared as described in the standard protocol from Bio-Rad(*110*). Compositions of buffers and solutions used can be seen in Table 3.

Table 3: Composition of buffers and solutions used for SDS-PAGE.

Buffer	Composition
Running buffer	3.0 g Tris base, 14.4 g glycine and 5 ml 20 % SDS diluted to 1 L with water
Sample buffer	2.0 ml water, 2.0 ml 1.0 M Tris/HCl pH 6.8, 4.0 ml glycerol, 2.0 ml 10 % SDS, 4 mg Coomassie Blue G-250 Before use 28 µl β-mercaptoethanol was mixed with 190 µl of sample buffer
Fixative solution	500 ml water, 400 ml 96% ethanol, 100 ml glacial acetic acid
Staining solution	900 ml water, 100 ml glacial acetic acid, 0.25g Coomassie blue G-250
De-staining solution	900 ml water, 100 ml glacial acetic acid

The marker contained 0.5 µl protein-standard (nr. 161-0304 from Bio-Rad) and 19.5 µl water. All samples were mixed with 5 µl sample buffer and heated at 80 °C for 10 min. The samples were loaded into the wells and the proteins were separated at 100 V for approx. 90 min with running buffer. After electrophoresis the gels were transferred to a fixing solution and soaked for 30 min. The gels were then stained for 1 hour in a staining solution, followed by overnight de-staining.

4.2.5.2 Activity assay

The activity measurements were performed on a Varian Cary Eclipse Fluorescence Spectrophotometer. The excitation wavelength was 300 nm and the emission was measured at 330 nm. Excitation and emission slits were 5 nm and the photomultiplicator tube voltage was 650 V. The measurements were done in 10.00 mm x 10.00 mm QS quartz cuvettes from Helma. The fluorescence spectrophotometer was equipped with a temperature-controlled-four-cuvette holder with stirring. All measurements were conducted at 15 °C.

For standard activity measurement, the TPH sample was thawed on ice and centrifuged for 15 min at 17000 x g in a Heraeus Pico centrifuge at 4 °C. The supernatant was collected, and the concentration was determined by at least 3 consecutive measurements using the NanoDrop spectrophotometer and the corresponding extinction coefficient at 280 nm. The sample was diluted to a concentration of 9.4 µM in the corresponding gel-filtration buffer and stored on ice.

The assay volume was 2500 µl and contained the components shown in Table 4. Buffer, premix, TPH sample, and MilliQ water was added to the cuvette containing a small magnetic stirring bar. The cuvette was placed in the sample holder, stirring was started, and the temperature was allowed to equilibrate for 2 min. The measurement was started, and after 30 s BH_4 was added to initiate the reaction.

Table 4: Concentration and volumes of stock solutions used in the activity assay.

Stock solution	Content	Concentration in stock solution	Volume added to assay	Concentration in assay
Buffer (room temperature)	HEPES/NaOH, pH 7,0 $(NH_4)_2SO_4$	125 mM 500 mM	1000 µL	50 mM 200 mM
Premix (freshly mixed)	Catalase (-20 ˚C)	1,25 g/L	50 µL	0,025 g/L
	$(NH_4)_2Fe(II)(SO_4)_2 \cdot 6H_2O$ (freshly prepared)	1,27 mM	50 µL	25 µM
	DTT (freshly prepared)	175 mM	100 µL	7 mM
	Tryptophan (-20 ˚C)	3,0 mM	50 µL	60 µM
TPH (-80 ˚C)	TPH diluted in purification buffer	9.4 µM	400 µL	1.5 µM
MilliQ water			800 µL	
BH_4 (-20 ˚C)		15 mM	50 µL	300 µM

The initial slope (int/min) was determined using the fluorescence spectrophotometer software. The software performs linear regression on an interval of the progression curve. The interval was selected manually and individually for each measurement, but stretched over minimum 0.04 min. The initial rate was converted to µM 5-OH-trp/min from standard curves produced in earlier projects(*9*). To calculate the specific activity, this value was divided by protein concentration.

When activity was measured during purification, the protein concentration was often low. In order to determine if any TPH was present in such low concentration samples, the temperature of the assay was occasionally increased. When determining specific activity, the measurements were always performed at 15 °C.

4.2.5.3 Stability studies by DSF

Differential scanning fluorimetry (DSF) can be used to determine the melting or unfolding temperature of a protein under different conditions. Here, it is used to study the stabilizing effect of different buffer compositions. All DSF measurements were performed at the Cassiopeia crystallization facility at MaxLab in Lund under supervision of Facility Manager Dr. Maria Håkansson. The results shown in this thesis are all from home-made non-commercial screens.

All screen solutions were prepared with concentrations 1.3 x the final concentration. They were transferred to a deep well block in volumes of 1.5 mL and frozen until use. At the crystallization facility, the deep well block was thawed and a Tecan pipetting robot was used to distribute the screen solutions to a white ABgene 96 well PCR plate (low profile, no skirts). 24 µL screen solution was added to each well. Protein solutions were made by diluting the concentrated protein to 300 µg/mL in 10 mM HEPES/NaOH, pH 7.0 (c*h*TPH2 and ct*h*TPH2) or 20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0 (c*h*TPH1 and *h*TPH1). This gives a protein concentration of 75 µg/mL in the final screen solution. SyPro orange was added in a volume ratio of 1:333, meaning that 1.5 µL SyPro orange was added to 500 µL protein solution. 8 µL protein solution with SyPro orange was added to each well in the PCR plate already containing 24 µL screen solution, to a total volume of 32 µL. The PCR plate was covered with clear PCR seal film and centrifuged at 200 x g for 1 min. It was then placed in an AH diagnostics Stratagene Mx3005p realtime PCR instrument. The fluorescence of SyPro orange was measured in the temperature interval 25-95 °C with an increase of 1 °C/min using an excitation wavelength of 492 nm and an emission wavelength of 610 nm. The melting temperature of the protein was determined as the maximum of the first derivative of the fluorescence curve.

The different screen conditions that were tested will not be presented here, but can be found in the results and discussion in section 4.3.2.3 and 4.3.3.3.

4.2.5.4 Mass spectrometry

All mass spectrometric measurements were performed on a LCT Premier mass spectrometer from Waters, Manchester, UK. The mass spectrometer is equipped with a Z-spray nano-flow ESI (electrospray ionization) source and an orthogonal acceleration time-of-flight (oa-TOF) mass analyzer(*111*).

For mass spectrometric analysis, protein samples were thawed, and exchanged into the volatile mass spectrometry buffer using Micro Bio-Spin 6 Chromatography columns from Bio-Rad. The mass spectrometry buffers used contain NH_4OAc in different concentrations (0.3-1.0 M) and some had pH adjusted. For denaturing conditions, 2-5 % formic acid was used. Some samples were filtrated using a 0.22 µm Millex-GV syringe filter. The protein concentration varied between 2-25 µM.

The instrument was calibrated with a solution of cesium iodide containing 100 mg/mL in 50 % 2-propanol. Typical instrument settings were: Sample cone 25 V, capillary voltage 1100-1800 V, cone gas-flow 20 L/h, source temperature 30 °C and ion guide 5-150 V. The function MaxEnt1 in the software suite MassLynx 4.1 was used to determine the mass of the detected ions.

4.2.5.5 Monodispersity test

The monodispersity of purification products were tested after concentrating and freezing, following the general method described here. The protein sample was thawed, and either centrifuged (2000 x g, 3 minutes) or filtrated. 250-500 µL sample was loaded on an analytical gel-filtration column corresponding to the gel-filtration column used in the last purification step (Superdex 75 10 300 GL or Superdex 200 10 300 GL). The column was equilibrated in the same buffer used in the last gel-filtration step in the purification. If several peaks were observed in the first run, the main peak was collected and loaded on the column again, to see if the sample remains monodisperse or if an equilibrium between different oligomer species is present.

4.3 Results and discussion

In this section, the results from the purification part of this project will be presented and discussed. First, the results from the purification and cleavage test of proteases will be given. This is followed by a presentation of the purification results for TPH1 variants and then TPH2 variants. In the end of this section an overall discussion is provided comparing the many purification and characterization results from different variants giving an overview of the data obtained in this project.

4.3.1 Purification of proteases and cleavage test

Both proteases were purified using a simple one-step purification method with affinity chromatography. The proteases were then exchanged into a storage buffer. If the proteases are stored in elution buffer, cleavage is less effective and requires large amount of protease (data not shown). Furthermore, the protease looses activity upon freezing if kept in elution buffer, however exchange into storage buffer eliminates this activity loss. The storage buffer resembles that from previous publications on 3CP(*112*, *113*), and includes DTT, EDTA and glycerol for stabilization.

4.3.1.1 Purification of GST-3CP

The procedure described in section 4.2.2.1 was followed and GST-3CP was collected from the elution peak of the affinity step. SDS-PAGE analysis shows that the solution contains GST-3CP, and some GST impurity (data not shown). Both can be removed from the target-protein solution after cleavage using affinity purification. The yields of GST-3CP were 5-8 mg/L culture, which is rather low compared to the yield of MBP-3CP (*vide infra*, see section 4.3.1.2).

4.3.1.1.1 Cleavage test

A cleavage test was performed in order to determine which ratio of protease to protein that is optimal and gives complete cleavage, as described in section 4.2.2.1.1. GST-c*h*TPH1 purified as described in section 4.2.3.1 was used as test protein. The cleavage test included different molar ratios of protease to protein, and the resulting SDS-PAGE can be seen in Figure 13. A complete cleavage is seen for all ratios from 1:1 to 1:7. For the ratio 1:15 there is a weak band representing GST-c*h*TPH1, indicating that cleavage was not complete. This band becomes even stronger for ratios 1:29 and 1:59. A ratio of protease to protein of 1:7 was chosen as optimal, since the cleavage is complete and it allows for the smallest consumption of protease. The latter is also an important parameter, especially since the yield of the protease purification is rather low.

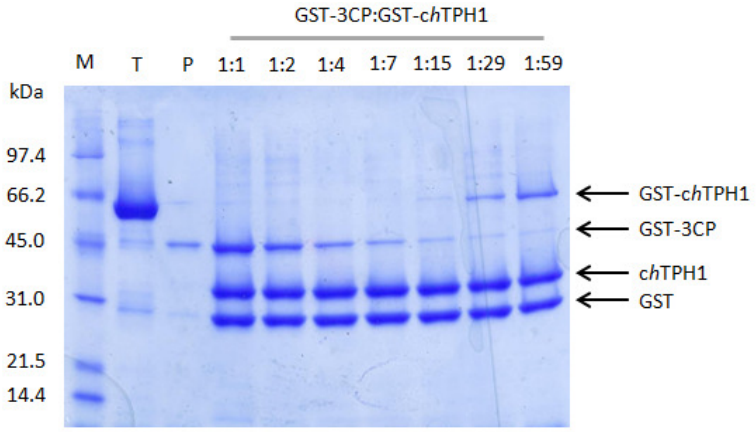


Figure 13: SDS-PAGE analysis of the cleavage test for GST-3CP. M: Molecular weight standard. T: Test protein GST-*ch*TPH1. P: Protease GST-3CP. The molar ratio between GST-3CP and GST-*ch*TPH1 is given above each lane. The concentration of GST-*ch*TPH1 and the total volume were constant in all samples.

4.3.1.2 Purification of MBP-3CP

The procedure described in section 4.2.2.2 was followed, and MBP-3CP was collected from the elution peak in the affinity purification. SDS-PAGE shows that the solution contains MBP-3CP (data not shown). The yields of purified MBP-3CP were 30-35 mg/L culture, which is much higher than for GST-3CP even considering the molecular weight difference (63 kDa for MBP-3CP vs. 47 kDa for GST-3CP).

4.3.1.2.1 Cleavage test

Again, a cleavage test was performed as described in section 4.2.2.2.1. Even though MBP-3CP and GST-3CP only differs in the affinity tag, there can be differences in cleavage capacity as the affinity tag may affect the protease e.g. by making dimers. Also, the test-protein in this cleavage test (MBP-*h*TPH1) contains the MBP-affinity tag plus a peptide linker. This linker might ease access to the recognition site compared to GST-fused proteins, and this might affect cleavage efficiency. The result of the cleavage test can be seen in Figure 14, and shows complete cleavage for ratios 1:3 and 1:7. Ratios of 1:50 and above lead to incomplete cleavage. Again, a ratio of 1:7 was chosen as the optimal condition for cleaving. The results show that there are no differences between GST-3CP and MBP-3CP in cleavage capacity.

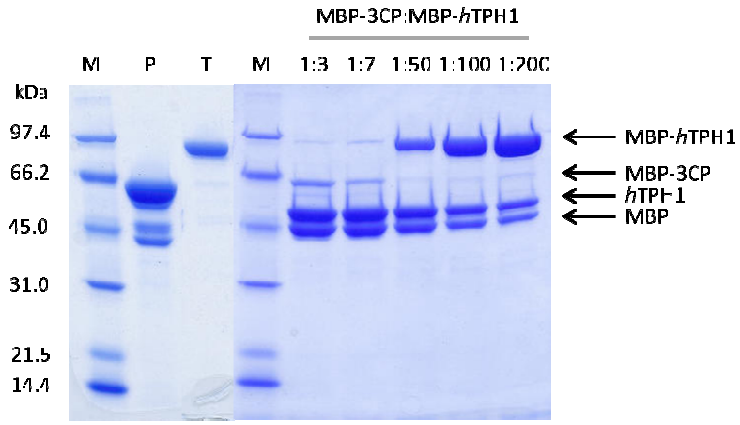


Figure 14: SDS-PAGE analysis of the cleavage test for MBP-3CP. M: Molecular weight standard. P: Protease MBP-3CP. T: Test protein MBP-*h*TPH1. The molar ratio between MBP-3CP and MBP-*h*TPH1 is given above each lane. The concentration of MBP-*h*TPH1 and the total volume were constant in all samples.

4.3.2 Purification of TPH1

In this section, the results from the development of different purification methods for TPH1 will be presented. First, the purification of GST-tagged *ch*TPH1 will be described, followed by purification of MBP-tagged *h*TPH1 and rc*h*TPH1. A section on stability studies of purified *ch*TPH1 and *h*TPH1 using DSF is also included.

4.3.2.1 Purification of GST-tagged chTPH1

A method for purification of *ch*TPH1 has been developed in earlier projects(*13*). This methods includes an anion-exchange step followed by a gel-filtration step, and can be performed in one day. However, a mass spectrometric analysis of the intact *ch*TPH1 product from this purification method showed significant impurities(*10*). Thus, a new method for purification was developed using GST as affinity tag.

The overall purification method consists of an affinity step followed by a gel-filtration step. Then the sample is cleaved overnight by GST-3CP and the product is again purified by an affinity step and a gelfiltration step. The first gel-filtration step is included to remove any gluthathione in the solution left over from the first affinity step, so it does not interfere with the second affinity step. The last gel-filtration step allows for buffer change and for checking monodispersity and size of the purified protein.

The chromatogram from the first affinity step in the purification of GST-*ch*TPH1 can be seen in Figure 15.

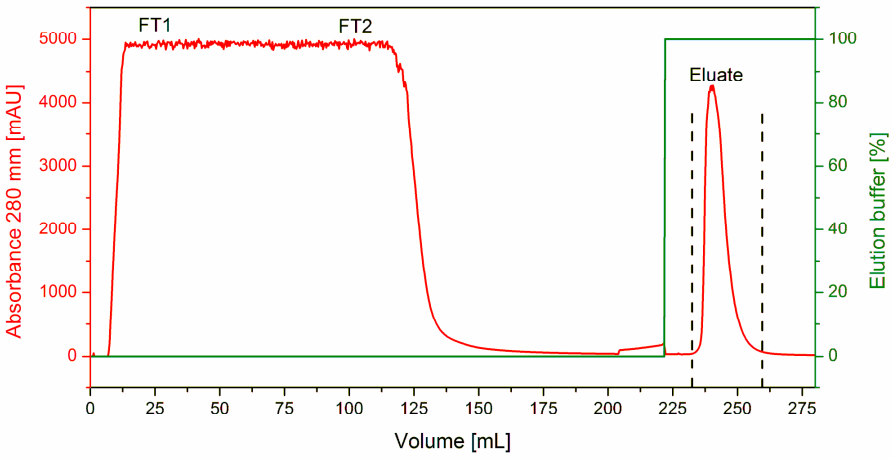


Figure 15: Chromatogram from the affinity purification of GST-*ch*TPH1 on a GST FF 16/10 column. FT1: First part of the flow-through. FT2: Last part of the flow-through.The fractions marked with dashed lines were collected for further purification.

GST-*ch*TPH1 was collected from the eluate, concentrated and loaded onto a gel-filtration column, resulting in the chromatogram shown in Figure 16. The major peak (P3) contains GST-*ch*TPH1. P1 and P2 also contain GST-*ch*TPH1 and must be oligomer forms of the protein. According to the calibration curve of the column and the elution volume, GST-*ch*TPH1 exists as a dimer or even trimer with an approx. size of 170 kDa (GST-*ch*TPH1: 62.5 kDa). As GST is known for dimerizing(*105*), GST-*ch*TPH1 most likely exists as a dimer.

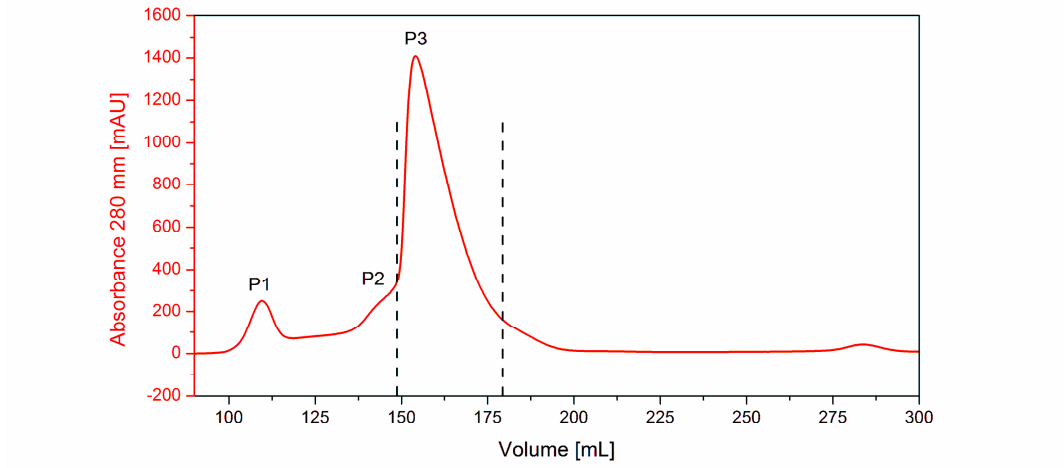


Figure 16: Chromatogram from the gel-filtration of GST-*ch*TPH1 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines were collected for cleavage and further purification.

The purified GST-*ch*TPH1 was cleaved overnight using GST-3CP. The cleaved sample was loaded on the affinity column again, and the flow-through containing *ch*TPH1 was collected, as shown in Figure 17. In this step the sample is loaded very slowly (1 mL/min) to allow for all GST species to bind to the column. Earlier purifications have shown that if the loading flow-rate is 2.5 mL/min, some GST does not bind to the column, and is thus found in the flow-through together with *ch*TPH1. GST and *ch*TPH1 cannot be separated in the last gel-filtration step due to their very similar sizes (36 kDa for *ch*TPH1 vs. 26 kDa for GST), and thus it is of great importance that the sample is loaded slowly. In the protocol for the GST column it is also recommended to keep the flow rate low during sample loading for maximum binding capacity, due to the relatively slow binding kinetics between GST and gluthathione. A lower flow rate thus ensures a longer residence time.

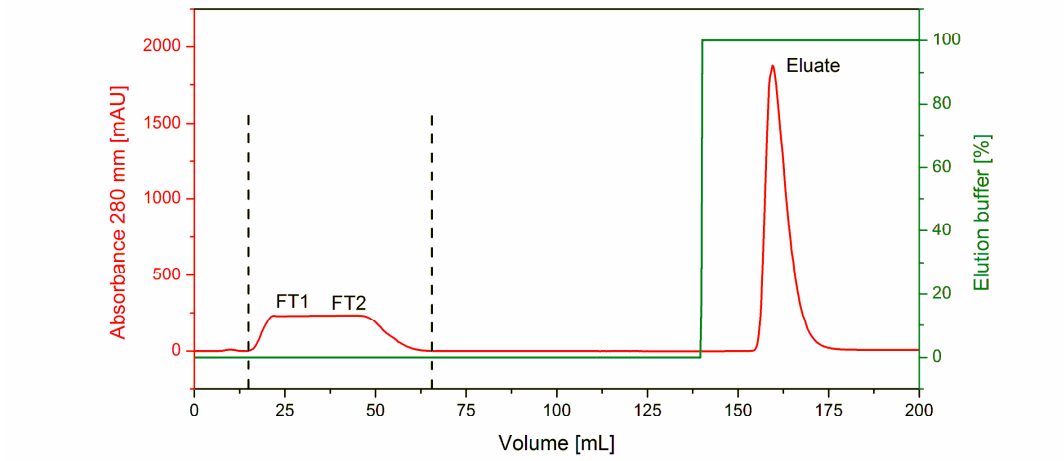


Figure 17: Chromatogram from the affinity purification of cleaved GST-*ch*TPH1 on a GST FF 16/10 column. FT1: First part of the flow-through. FT2: Last part of the flow-through. The fractions marked with dashed lines were collected for further purification.

The flow-through of the last affinity step is collected, concentrated and loaded on a gel-filtration column. As *ch*TPH1 is expected to be monomeric, a superdex 75 column is used instead of superdex 200 as in the first gel-filtration step. The chromatogram in Figure 18 shows only one peak (P4), which contains *ch*TPH1.

The protein solution was thus pure already before the last gel-filtration step, which could be omitted. However, if buffer exchange is needed, the gel-filtration step can be included. Also, the chromatogram confirms that c*h*TPH1 exists as a monomer (approx. 40 kDa according to elution volume) and is monodisperse before freezing.

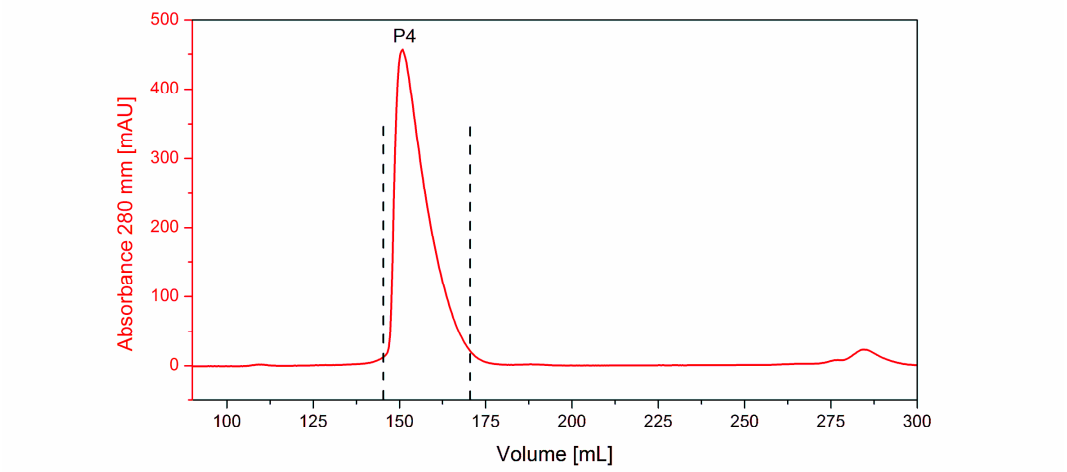


Figure 18: Chromatogram from the gel-filtration of cleaved GST-c*h*TPH1 on a HiLoad Superdex 75 26/60 prep grade column. The fractions marked with dashed lines were collected as the final product.

The SDS-PAGE analysis of the entire purification is shown in Figure 19. GST-c*h*TPH1 is partly soluble as a relatively large amount is found in the pellet after sonication. The pellet has been resuspended in the same volume of buffer as the raw extract, so the bands can be directly compared. However, approx. 60 % of the expressed GST-c*h*TPH1 is soluble and found in the supernatant.

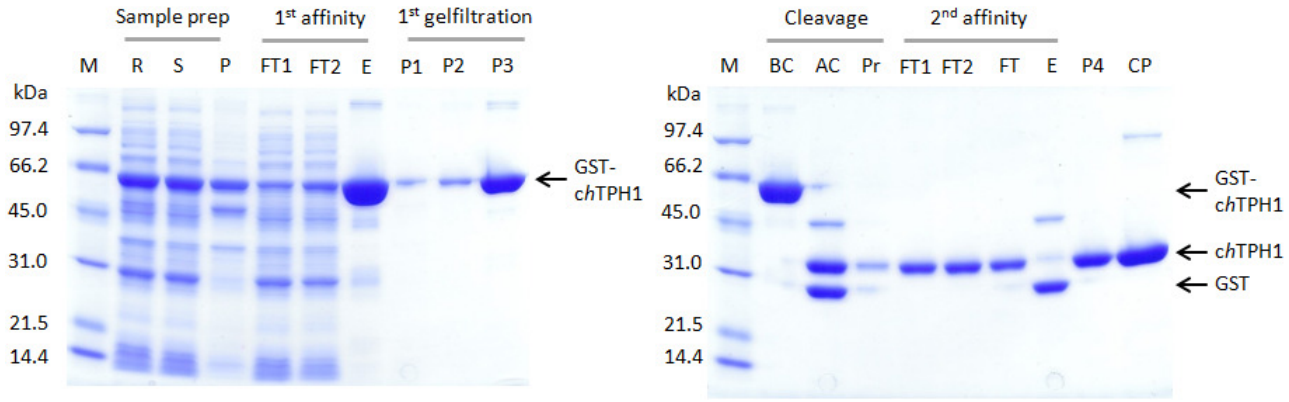


Figure 19: SDS-PAGE analysis of the purification of GST-tagged c*h*TPH1. M: Molecular weight standard. R: Raw extract. S: Supernatant. P: Resuspended pellet. FT1: First part of the flow-through from the GST-column during sample loading. FT2: Last part of the flow-through. E: Eluate. P1-P3: Peaks from the first gel-filtration. BC: Before cleavage. AC: After cleavage. Pr: Precipitate observed after cleavage. FT: Entire flow-through mixed. P4: Peak from the second gel-filtration.

In the first affinity step, minor amounts of GST-c*h*TPH1 are found in both the first and last part of the flow-through. This suggests that the loading speed (2.5 mL/min) is too fast to ensure a complete binding of GST-c*h*TPH1. If the column had been overloaded, GST-c*h*TPH1 would only be seen in the last part of the flow-through. The loading speed of this first affinity step was kept at 2.5 ml/min despite the loss of sample, as the step would otherwise be very time-consuming (the sample size is 120 mL). The loss of target protein in

this first step does not affect the purity of the final sample but merely the yield of the purification. In the eluate from the first affinity column, most GST-c*h*TPH1 is found. All three peaks (P1, P2, and P3) from the first gel-filtration contain GST-c*h*TPH1, and P1 and P2 must thus represent oligomer species (soluble aggregates). The sample of GST-c*h*TPH1 from the first gel-filtration appears pure, and can be completely cleaved. In some purifications, precipitation was observed after cleavage. This precipitate apparently contains c*h*TPH1, and thus the precipitation should lower the final yield. The flow-through from the second affinity step contains pure c*h*TPH1, while GST and GST-3CP is found in the eluate. P4 from the last gel-filtration contains pure c*h*TPH1.

The purity of the final product was also investigated with SDS-PAGE by loading different volumes of the product. Figure 20 shows, that the c*h*TPH1 solution is completely pure, as no impurities are seen on the gel.

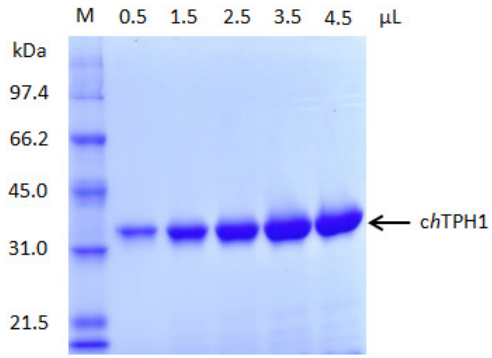


Figure 20: SDS-PAGE analysis of the final product of purified c*h*TPH1. M: Molecular weight standard. Numbers: Volume in µL loaded on the gel.

The purity was furthermore checked by mass spectrometry, and the mass spectrum of intact c*h*TPH1 is seen in Figure 21. The calculated mass of the only species present is 36540.9 Da. This mass corresponds very well with the theoretical mass of holo-c*h*TPH1, which is 36530.6 Da. It can once again be concluded that the sample is pure and contains c*h*TPH1. Addition of EDTA to the sample gave a decrease in the mass of 57.9 Da corresponding to the removal of Fe(III) with a theoretical mass of 52.8 Da (data not shown).

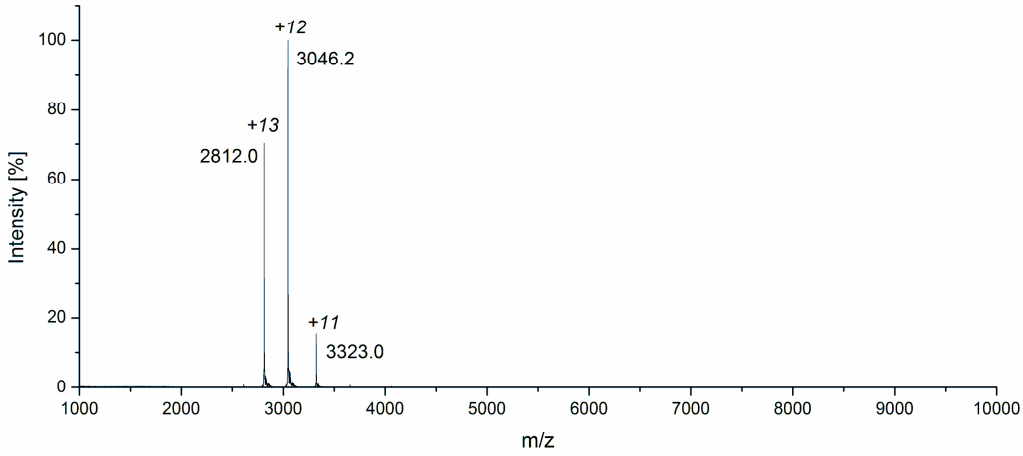


Figure 21: Mass spectrum of 4 µM c*h*TPH1 in 600 mM NH_4Ac/NH_3, pH 7.5. The calculated mass is 36540.9 Da.

To further characterize purified c*h*TPH1, it was denatured using 2 % formic acid. The resulting mass spectrum is seen in Figure 22. Again, only one species is present with a mass of 36479.4 Da. The theoretical mass of apo-c*h*TPH1 is 36477.8 Da, so again the masses match very well.

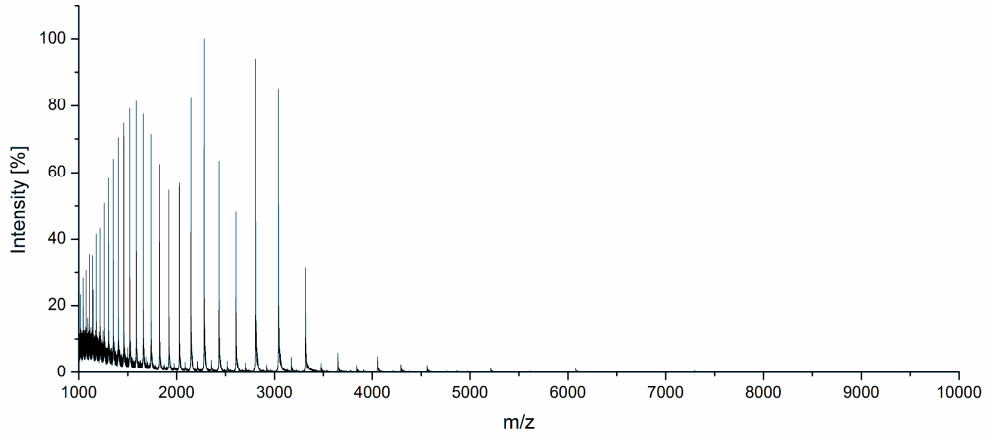


Figure 22: Mass spectrum of 4 µM denatured c*h*TPH1 in 30 mM NH_4OAc incubated with 2 % formic acid. The calculated mass is 36479.4 Da.

Finally, the purified product was tested in an activity assay (see section 4.4). Here, the specific activity is given as µmol 5-OH-trp/µmol TPH/min in units of min^{-1} as this allows for a direct comparison between different TPH variants. c*h*TPH1 showed a specific activity of 92 ± 2.2 min^{-1} which is in the same range as previously determined(*13*).

It can be concluded that the purification procedure gives pure and active c*h*TPH1, and that the previous impurity is not present.

The overall yield of the purification was in the range 10-11 mg/L culture. UV-Vis measurements were performed after each step in order to identify any critical steps in the purification, see Table 5. When calculating the yield, all absorbance at 280 nm is assigned to TPH, and thus removal of impurities will also result in a decrease in yield. The product from the first affinity purification is considered 100 %, as it is not possible to determine the amount of TPH in the raw extract by UV-Vis. From the SDS-PAGE analysis it is known that some GST-c*h*TPH1 is lost in this first affinity step. This loss is most likely due to the high flow-rate of 2.5 ml/min and not because the column is overloaded, as the yield of the first step is significantly lower than the binding capacity of the column (approx. 6-7 µmol).

Table 5: Protein yield after each step in the purification of GST-tagged c*h*TPH1. The yield was calculated from the absorption at 280 nm assuming that the samples contained only GST-c*h*TPH1 or c*h*TPH1 (after cleavage). The yield in % is given assuming that the first affinity step corresponds to 100 %.

	GST-c*h*TPH1	
	Yield	
Step	**[µmol]**	**[%]**
1st affinity step	2.4	100
1st gel-filtration step	1.0	43
2nd affinity step	1.0	42
2nd gel filtration	0.6	25

Approx. half of the protein is lost in the first gel-filtration. The cleavage and second affinity step is very effective and retains almost all c*h*TPH1, despite the observed precipitation. The last gel-filtration step again has some loss of protein most likely due to concentration of the sample by ultrafiltration prior to sample loading. The loss of protein in the last gel-filtration step suggests that if the yield of the process should be optimized, the last purification step should be omitted, as it does not increase the purity of the sample.

4.3.2.1.1 Monodispersity

For crystallization and most other characterization techniques, a monodisperse sample usually gives the best results(*95*). In the crystallization setup process, the purified protein is thawed and directly used. To test the monodispersity of the frozen protein solution, the final product from the purification was thawed and loaded onto an analytical gel-filtration column. The resulting chromatograms are shown in Figure 23. The first run reveals that the thawed c*h*TPH1 solution is not completely monodisperse, as a small peak is seen in front on the main peak. The small peak most likely represents aggregated c*h*TPH1. The main peak from the first run was collected and loaded on the analytical gel-filtration column again. This time there is only one visible peak in the chromatogram. From the chromatogram of the last gel-filtration in the purification shown in Figure 18 it is known that purified c*h*TPH1 is monodisperse right after purification. Thus, it can be concluded that even though c*h*TPH1 is monodisperse right after purification, it becomes slightly polydisperse upon freezing and thawing. The monodispersity of the sample can be regained by gel-filtration.

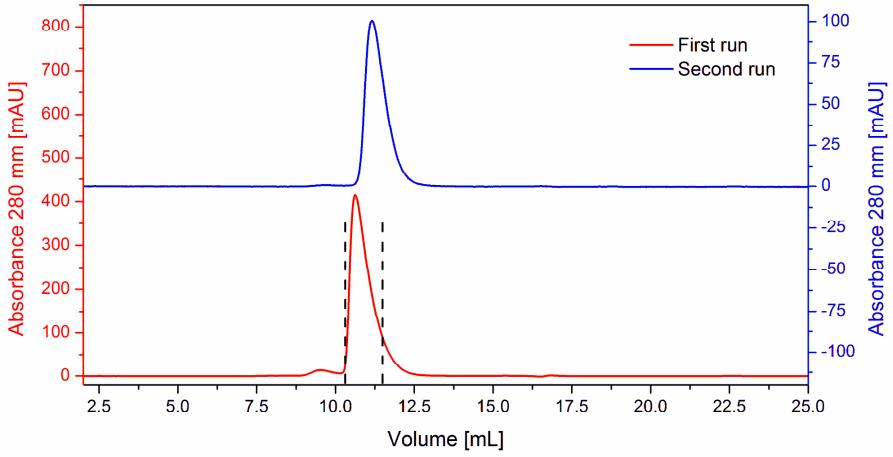


Figure 23: Chromatograms from the analytical gel-filtrations of thawed c*h*TPH1 on a Superdex 75 10 300 GL column. Red: The first run of thawed c*h*TPH1. Fractions marked with dashed lines were collected and loaded again for the second run. Blue: The second run of c*h*TPH1 collected from the main peak in the first run. The slightly shifted positions of the two peaks arise from different loading volumes (250 µL in the first run, 500 µL in the second).

4.3.2.2 Purification of MBP-tagged h*TPH1*

The successful purification of GST-c*h*TPH1 inspired to also test this new method involving affinity purification with full-length *h*TPH1. As GST was not suitable as affinity tag for the tetrameric *h*TPH1 (*vide infra*), MBP was instead used as the preferred tag. The purification method of *h*TPH1 presented here was developed through several stages. First, the method only contained an affinity step after the cleavage, but SDS-PAGE analysis revealed that MBP impurities were still found in the product. Thus, an extra gel-filtration step was added at the end of the purification.

First, the solubility of GST-*h*TPH1 was investigated in 20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0. GST was found as a suitable affinity-partner in purification of c*h*TPH1, thus it was also tested with *h*TPH1. The main difference between the two is that *h*TPH1 exists as a tetramer, whereas c*h*TPH1 is a monomer. As seen in Figure 24, most GST-*h*TPH1 is insoluble and found in the pellet. The insolubility of GST-*h*TPH1 is most likely due to the ability of GST to dimerize, which makes GST-tags unsuitable for the isolation of oligomeric proteins like *h*TPH1(*114*). The dimerization of GST could result in large insoluble aggregates, when it is fused to *h*TPH1 which forms tetramers. In Figure 24, the solubility test of non-affinity tagged *h*TPH1 performed in reference (*94*) is also shown. Here it is seen that non-tagged *h*TPH1 is very insoluble, with only about 10 % of the expressed protein present in the soluble fraction. This is very similar to GST-*h*TPH1 and it can thus be speculated that the GST-tag does not alter the solubility of *h*TPH1, but simply retains the original solubility of *h*TPH1. It should be noted that the solubility test of non affinity-tagged *h*TPH1 was performed in a buffer with no salt, while the solubility test of GST-*h*TPH1 was performed in a buffer with 300 mM $(NH_4)_2SO_4$.

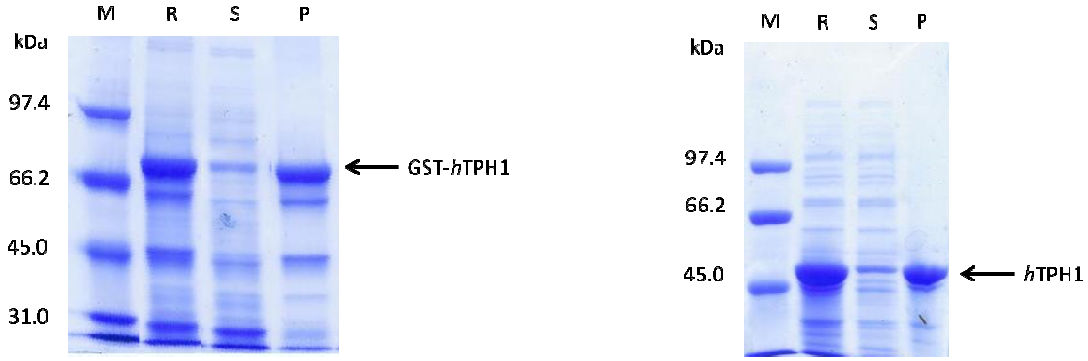


Figure 24: SDS-PAGE analysis of the solubility test of GST-*h*TPH1 and *h*TPH1. M: Molecular weight standard. R: Raw extract. S: Supernatant. P: Resuspended pellet. Left: Solubility test of GST-*h*TPH1 in 20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0. Right: Solubility test of *h*TPH1 in 20 mM Tris/H_2SO_4, pH 8.5. Adapted from (*94*)**.**

Instead, *h*TPH1 was fused to MBP, which does not oligomerize. MBP is a commonly used affinity tag, which is known for enhancing the solubility of the fusion-partner(*115*) if attached at the N-terminal(*116*). It has also been used as affinity tag in the published purification methods for *h*TPH1 and *h*TPH2(*51*). As seen from Figure 25, most of the expressed MBP-*h*TPH1 is soluble and found in the supernatant. This confirms that MBP is better suited as affinity tag for *h*TPH1 than GST.

When comparing the soluble (supernantant, S) and insoluble (pellet, P) fractions, approx. 90 % MBP-*h*TPH1 is present in the soluble fraction. This is in contrast to earlier expressions of MBP-*h*TPH1 by McKinney et al.(*98*), where half was found in the insoluble fraction. The main difference between the two expression systems is the induction temperature. In this project MBP-*h*TPH1 is expressed at 20 °C, while McKinney et al uses temperatures of 28 and 37 °C, respectively. It is well known that expression at lower temperatures can increase the amount of soluble protein in *E. coli*(*117*).

The effect of pH and salt on the solubility of MBP-*h*TPH1 were tested and the result is seen in Figure 25. Neither pH nor salt seems to have a significant effect on the solubility of MBP-*h*TPH1. Earlier studies have shown that GST-rhTPH1 requires high concentrations of salt during purification to keep it stable(*10*). Thus, the high concentration of $(NH_4)_2SO_4$ was chosen for the purification of *h*TPH1 despite the lack of effect on the solubility of MBP-*h*TPH1. The high salt concentration might affect the stability of TPH after cleavage

from the MBP affinity-tag. The pH of the buffer was chosen to 8.0. The theoretical pI of *h*TPH1 is calculated to 6.8(*109*), and 8.0 is well above this, theoretically ensuring a more soluble protein.

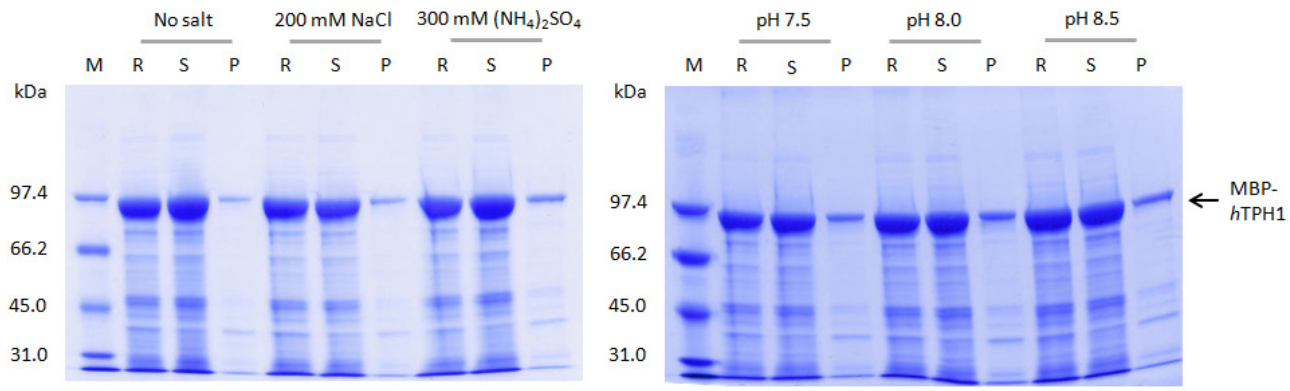


Figure 25: SDS-PAGE analysis of the solubility test of MBP-*h*TPH1. Left: Salt solubility test. In all samples the background buffer was 20 mM Tris/H_2SO_4, pH 8.0. Right: pH solubility test. In all samples the background buffer was 20 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$. M: Molecular weight standard. R: Raw extract. S: Supernatant. P: Resuspended pellet.

MBP-*h*TPH1 was purified according to the procedure described in section 4.2.3.2. It was first attempted to leave out several steps in the process to get as simple a procedure as possible. However, most steps turned out to be necessary to avoid impurities in the final product. The gel-filtration after the first affinity step was included to remove maltose from the solution, which would otherwise disturb the second affinity step. Furthermore, soluble aggregates of MBP-*h*TPH1 were observed, which could not be effectively cleaved, and these could also be discarded in the first gel-filtration step to minimize consumption of protease. The last gel-filtration step was included as the second affinity step did not completely remove all MBP species from the solution. The chromatogram from the first affinity purification is found in Figure 26. MBP-*h*TPH1 is found in the eluate, and is collected for further purification.

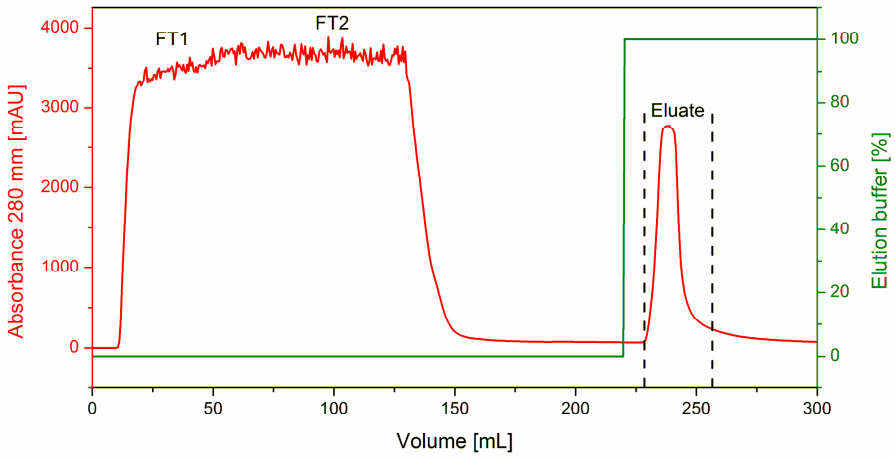


Figure 26: Chromatogram from the affinity purification of MBP-*h*TPH1 on a 20 mL Dextrin Sepharose HP 16/10 MBP column. FT1: First part of the flow-through. FT2: Last part of the flow-through. The fractions marked with dashed lines were collected for further purification.

The product from the first affinity purification was concentrated and loaded on a gel-filtration column. The resulting chromatogram is seen in Figure 27. Both peaks (P1 and P2) contain MBP-*h*TPH1. According to the

calibration curve of the column, P2 corresponds to tetrameric MBP-*h*TPH1 and this peak is selected for further purification. The first peak (P1) elutes in the dead volume of the column and contains MBP-*h*TPH1, which is also active (data not shown). Thus, this peak might represent soluble aggregates of MBP-*h*TPH1. Attempts to minimize this peak by testing different buffers failed. SDS-PAGE analysis shows that MBP-*h*TPH1 from this peak cannot be completely cleaved (data not shown), which supports the hypotheses that it contains large aggregates, where the protease might not be able to reach the inside to cleave effectively.

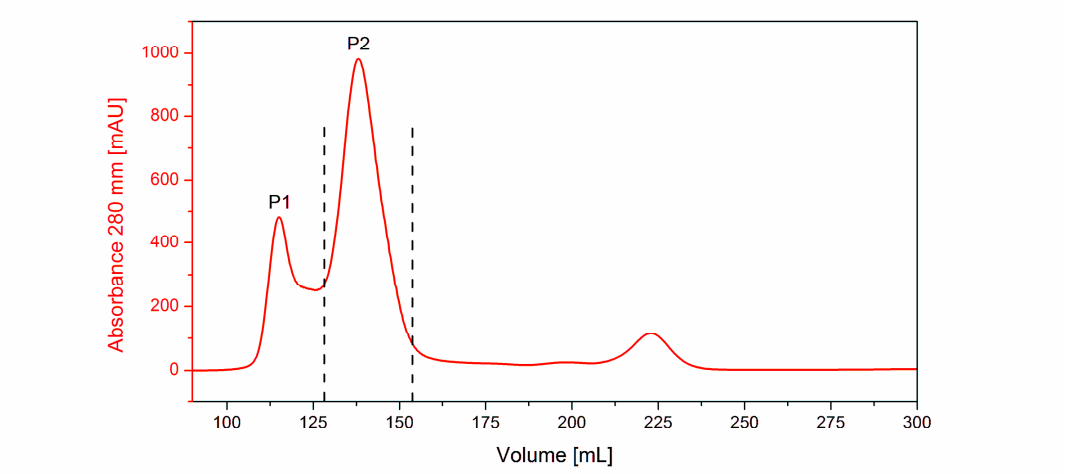


Figure 27: Chromatogram from the gel-filtration of MBP-*h*TPH1 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines were collected for cleavage and further purification.

SDS-PAGE analysis for the purification of MBP-*h*TPH1 up until cleavage is shown in Figure 28. It is seen that most MBP-*h*TPH1 binds to the affinity column and can be eluted. However, a small part is found in the flow-through, especially in the last part of the sample loading, indicating that the binding capacity of the column is exceeded. The eluate from the first affinity purification contains almost exclusively MBP-*h*TPH1. The SDS-PAGE analysis of the first gel-filtration shows that both peaks (P1 and P2) and the valley between them contain MBP-*h*TPH1.

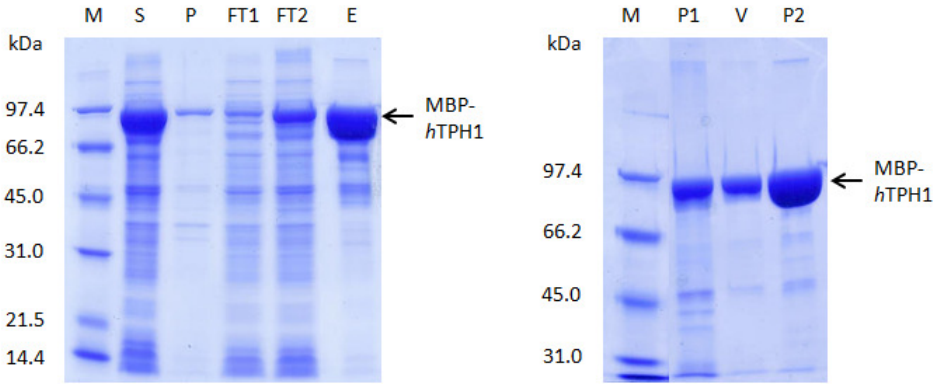


Figure 28: SDS-PAGE analysis of first part of the purification of MBP-*h*TPH1. Left: Sample preparation and first affinity step. M: Molecular weight standard. S: Supernatant. P: Resuspended pellet. FT1: First part of the flow-through during sample loading on the MBP affinity column. FT2: Last part of the flow-through. E: Eluate from the affinity column. Right: Gel-filtration. P1-P2: Peaks from first gel-filtration shown in Figure 27. V: Valley between P1 and P2.

The product from the gel-filtration containing MBP-*h*TPH1 was cleaved with MBP-3CP overnight, and loaded on the MBP column again. The chromatogram of this second affinity step can be seen in Figure 29. *h*TPH1 can be found in the flow-through, while most MBP and MBP-3CP binds to the column and is found in the eluate.

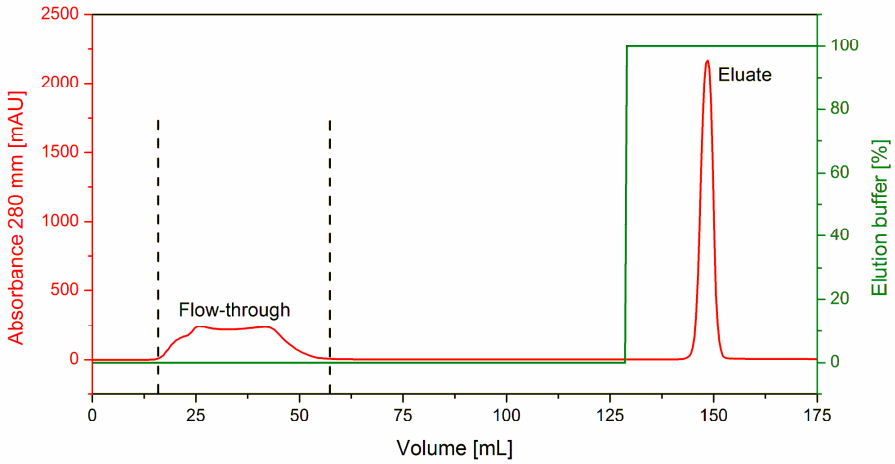


Figure 29: Chromatogram from the affinity purification of cleaved MBP-*h*TPH1 on a 20 mL Dextrin Sepharose HP 16/10 MBP column. The fractions marked with dashed lines were collected for further purification.

The flow-through containing *h*TPH1 is concentrated and loaded onto the gel-filtration column for the last purification step. The chromatogram is seen in Figure 30. The main peka P4 contains *h*TPH1, and according to the calibration curve the size corresponds to a tetramer as expected.

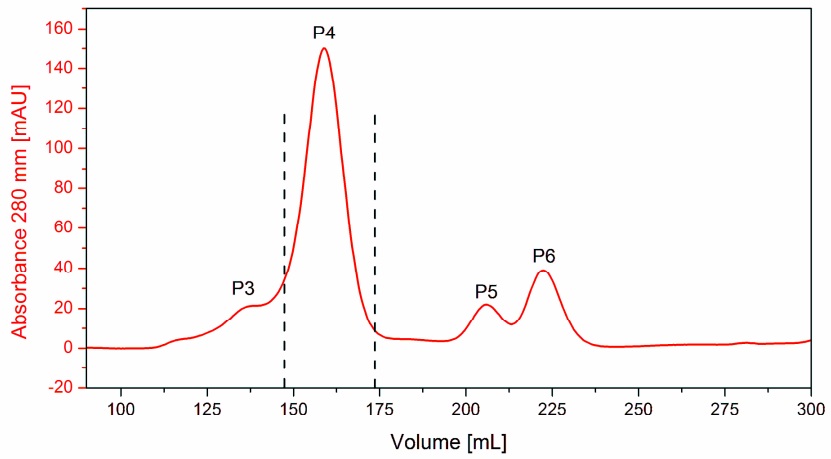


Figure 30: Chromatogram from the gel-filtration of cleaved MBP-*h*TPH1 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines were collected as the final product.

The SDS-PAGE analysis of the last part of the purification of *h*TPH1 is seen in Figure 31, and shows that the sample from the first gel-filtration containing MBP-*h*TPH1 is completely cleaved after 16 hours. Furthermore, the cleavage process seems rather fast, as it is almost complete after only 10 minutes. The flow-through from the second affinity step contains *h*TPH1 mixed with impurities of MBP and MBP-3CP. The last affinity step is thus not effective enough to remove all MBP species, all though a large part of the impurities have been removed. The SDS-PAGE analysis of the peaks from the last gel-filtration shows that

P4 contains *h*TPH1, and so does the shoulder on the left called P3. P5 contains MBP-3CP and P6 contains MBP, which have not been removed in the affinity step. The purity of the final product is confirmed by loading different volumes of the final product. No impurities are seen.

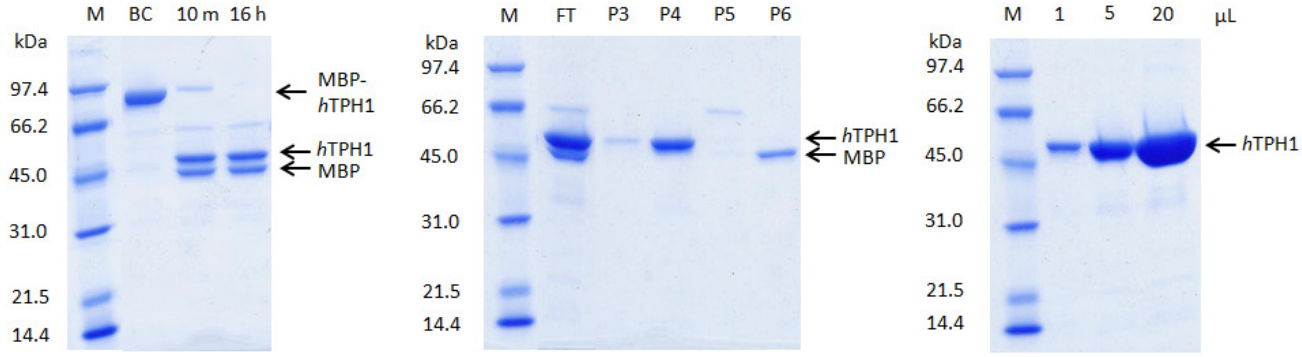


Figure 31: SDS-PAGE analysis of the second part of the purification of MBP-tagged *h*TPH1. Left: The cleavage of MBP-*h*TPH1 using MBP-3CP. M: Molecular weight standard. BC: Product from the first gel filtration before cleavage. 10 m: 10 minutes after addition of protease. 16 h: 16 hours after addition of protease. Center: Second affinity and gel-filtration steps. FT: Flow-through from the second affinity step. P3-P6: Peaks from the last gel-filtration. Right: The final product of *h*TPH1. Numbers: µL loaded on the gel.

To confirm the purity and identity of the final product, mass spectrometric measurements of the intact protein were made as described in section 4.2.5.4. The identification of large intact protein complexes such as tetrameric *h*TPH1 is a relatively new achievement, which represents state-of-the-art in protein mass spectrometry(*118*). The m/z spectrum obtained from 7 µM *h*TPH1 in 1 M NH_4OAc is shown in Figure 32. It shows only one major species, which is assigned to tetramer *h*TPH1. The calculated mass is 204929 Da. This corresponds well with the theoretical mass of tetramer *h*TPH1 including Fe(III) which is 204820 Da. It can thus be concluded that the final product from the purification contains pure tetrameric *h*TPH1.

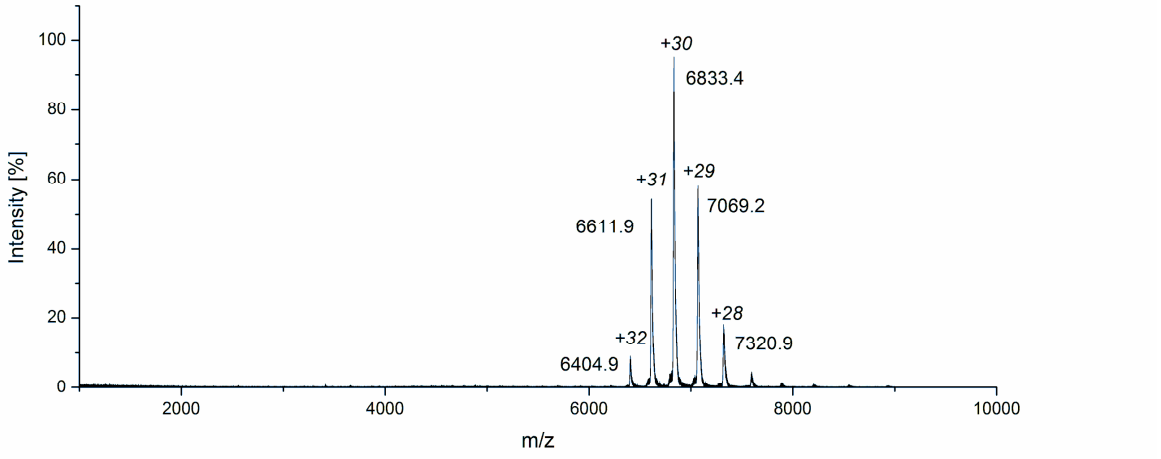


Figure 32: Mass spectrum of 7 µM *h*TPH1 in 1 M NH_4OAc. The calculated mass is 204929 Da.

Further investigations of the final product were performed by analyzing a denatured sample of *h*TPH1 also with mass spectrometry. Using 5 (v/v)% formic acid *h*TPH1 was partly unfolded, and the resulting mass spectrum in Figure 33 nicely shows the breakdown of the tetramer to dimer, monomer and unfolded (denatured) monomer. Mass analysis on the unfolded monomer gave a calculated mass of 51156 Da. This

mass is very similar to the theoretical mass of apo *h*TPH1 monomer of 51152 Da. This confirms once again that the sample contains *h*TPH1.

The final product of the purification was also tested in the activity assay, and purified *h*TPH1 showed significant activity for hydroxylation of tryptophan with a specific activity of 164 ± 3.1 min^{-1}.

It can be concluded that the purification method developed for *h*TPH1 gives pure, soluble and active *h*TPH1.

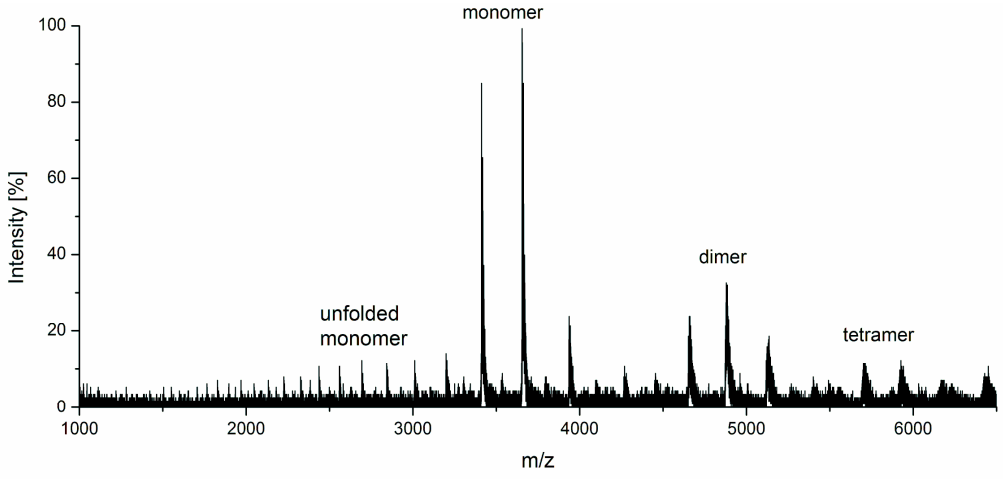


Figure 33: Mass spectrum of 7 µM *h*TPH1 in 0.33 M NH_4OAc and 5(v/v) % formic acid. The calculated mass of the unfolded monomer was 51156 Da.

The overall yield from the purification was in the range 16-23 mg/L culture. Furthermore, the yield of protein after each purification step was calculated, see Table 6. The yield is shown for both a standard purification, and for a purification with low salt concentration in all buffers (50 mM $(NH_4)_2SO_4$). The yield from the first affinity step was considered 100 %. From SDS-PAGE it is known that some MBP-*h*TPH1 is lost in the last part of the flow-through, probably due to the limited binding capacity of the MBP column. Comparison of the yield after the first affinity step (2.4 and 3.3 µmol) with the expected binding capacity of the MBP column (approx. 2.5 µmol depending on the protein) supports this hypothesis. However, if the flow-through is loaded on the affinity column again, almost no MBP-*h*TPH1 binds to the column, and 90 % is found in the flow-through again. The reason for this is unknown, but the phenomenon was also observed in purification of all other MBP-tagged TPH variants (*vide infra*).

Table 6: Protein yield after each step in the purification of MBP-*h*TPH1 with high and low salt concentrations (300 mM and 50 mM $(NH_4)_2SO_4$, respectively). The yield was calculated from the absorption at 280 nm assuming that the samples contained only MBP-*h*TPH1 or *h*TPH1 (after cleavage). The yield in % is given assuming that the first affinity step corresponds to 100 %.

	MBP-*h*TPH1 300 mM $(NH_4)_2SO_4$		MBP-*h*TPH1 50 mM $(NH_4)_2SO_4$	
	Yield		Yield	
Step	**[µmol]**	**[%]**	**[µmol]**	**[%]**
1st affinity step	2.4	100	3.3	100
1st gel-filtration step	1.1	48	1.8	28
2nd affinity step	1.2	52	1.6	27
2nd gel filtration	0.6	26	0.4	12

Appproximately half of MBP-*h*TPH1 is lost in the first gel-filtration step. This is most likely due to the aggregated species eluting in P1 in Figure 27. The increase in yield from the first gel-filtration to the second affinity step most likely originates from the presence of impurities in the flow-through of the affinity step after cleavage, which contributes to the absorption. The second affinity step is not completely effective, and some MBP and MBP-3CP impurities are still found in the flow-through. In the last gel-filtration almost half of *h*TPH1 is lost again. The decrease in yield could be caused by a combination of impurities that are removed and *h*TPH1 that precipitates or is lost during concentration.

The purification method developed here is very similar to the one developed by McKinney et al(*98*). It differs mainly in the buffers used. McKinney et al uses 20 mM Na–HEPES buffer, pH 7.0, 20 % glycerol, 1mM PMSF (serine protease inhibitor) for the affinity step, and 15 mM Na–HEPES, pH 7.4, 200 mM NaCl for the gel-filtration step of MBP-*h*TPH1. They also observe a high molecular weight peak in the gel-filtration containing MBP-*h*TPH1. They cannot detect any activity in the peak and assigns it to aggregated protein. Upon cleavage using enteropeptidase, TPH1 precipitated and was identified in the insoluble fraction by SDS-PAGE. Since the main difference between the purification method developed here and the one developed by McKinney et al. is the salt concentration in the buffer system, and since earlier studies have shown that GST-r*h*TPH1 requires high salt concentrations to be stable, it is reasonable to assume that *h*TPH1 also requires high salt concentrations for solubility. To test this hypotheses, a purification of MBP-*h*TPH1 was performed using only 50 mM $(NH_4)_2SO_4$ in the buffer. As expected, a lower yield of the purification was seen, despite a higher initial yield from the first affinity step (see Table 6). However, cleavage does not seem to be the critical step, as almost no *h*TPH1 is lost between the first gel-filtration and the second affinity step, where the cleavage takes place. The most significant difference between the purifications was the visible precipitation in many fractions when using low salt concentration. Also, the protein could only be concentrated to 50 µM in low salt, while a concentration of 150 µM was achieved in high salt. These results suggest that even though a high salt concentration is not absolutely required for purification of *h*TPH1, it helps to stabilize *h*TPH1 and increases the yield.

The purification method of *h*TPH1 can be optimized by excluding the second affinity step, and only performing a gel-filtration step after cleavage. The chromatogram in Figure 30 confirms that both MBP and MBP-3CP can be separated from *h*TPH1 using size-exclusion, so that the second affinity step is unnecessary. Also, it seems that the first affinity step is the limiting factor for the final yield. Thus, dividing the supernatant in two portions, and loading them in separate runs, could possible increase the yield of this first step. However, it would increase the time spent on purification significantly. Furthermore, the purification procedure can be shortened significantly by using on-column cleavage, which would shorten the purification procedure to just two steps that can be performed in one day.

4.3.2.2.1 Monodispersity

The chromatogram for the monodispersity test showed several shoulders besides the main peak indicating that the sample is not completely monodisperse, see Figure 34. The main peak was collected and loaded on the gel-filtration column again, in order to see if the monodispersity could be restored. The second run showed almost no shoulders and the sample can be considered monodisperse. It is not known whether the polydispersity is caused by the freezing/thawing process, the concentrating or simply by time. For crystallization purposes, the thawed protein solution should be loaded on a gel-filtration column before use or the drops should be set up directly after purification.

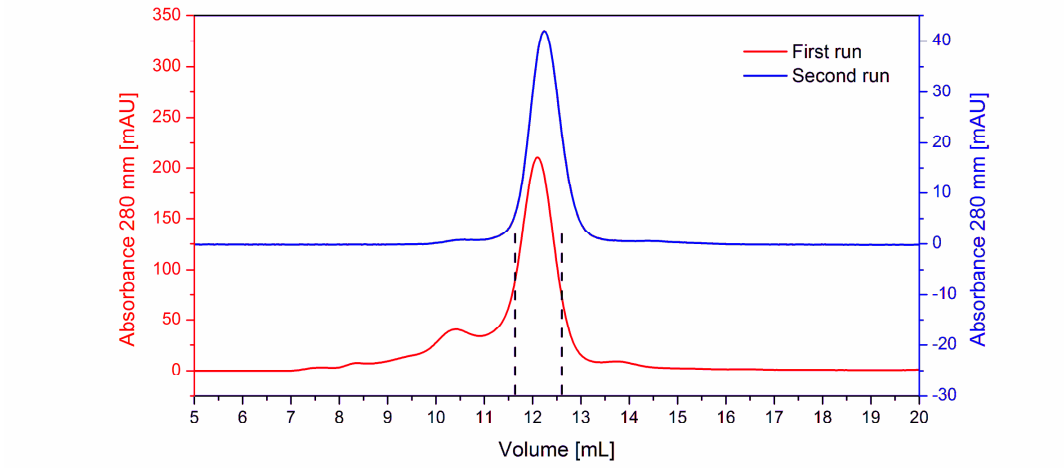


Figure 34: Chromatograms from the analytical gel-filtrations of thawed *h*TPH1 on a superdex 200 10 300 GL column. Red: The first run of thawed *h*TPH1. Fractions marked with dashed lines were collected and loaded again for the second run. Blue: The second run of *h*TPH1 collected from the main peak in the first run.

*4.3.2.3 Stability of ch*TPH1 *and* h*TPH1*

*ch*TPH1 and *h*TPH1 were studied by DSF to investigate the stability and effect of different buffer compositions. DSF can detect the temperature at which the protein unfolds under different conditions testing the thermal stability of the protein(*119*). The unfolding is measured by an increase in fluorescence of a dye with affinity for hydrophobic parts of the protein, which are exposed as the protein unfolds. The fluorescent dyes that can be used for DSF are highly fluorescent in non-polar environment, such as the hydrophobic sites on unfolded proteins, compared to aqueous solution where the fluorescence is quenched. A typical DSF result with schematic drawings of the processes involved is shown in Figure 35, left. When the protein is folded, fluorescence is low, but as the protein unfolds fluorescence increases. After unfolding, the proteins will aggregate and/or precipitate and the fluorescence will decrease again. The unfolding or melting temperature is determined as the maximum point of the first derivative of the data. DSF only requires very low concentrations and small amounts of protein to test many different conditions.

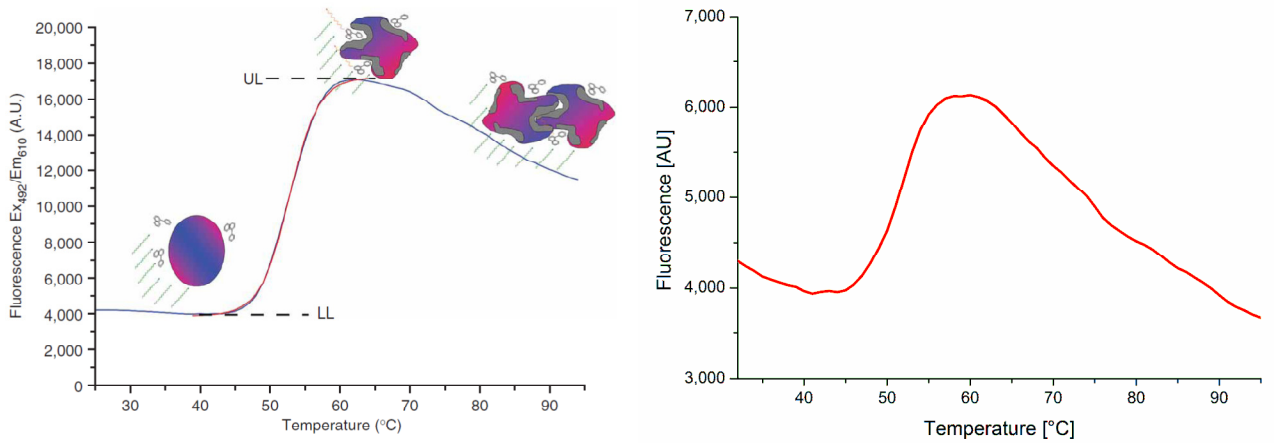


Figure 35: Left: A typical DSF result recording fluorescence as a function of temperature along with schematic drawings of the processes involved. UL and LL: Upper and lower level of fluorescence. Adapted with permission from (*119*). Copyright 2007 Nature Publishing Group. Right: DSF result for *h*TPH1 in the reference buffer 100 mM Tris/H_2SO_4, 300 mM $(NH_4)_2SO_4$, pH 8.0.

Fluorescence intensity as a function of temperature for *h*TPH1 is shown in Figure 35, right. It resembles that of a typical DSF curve in Figure 35, left, and shows only one major transition. Multi-domain proteins can show several unfolding transitions at different temperatures, as one domain can unfold without affecting the other domains, but this is not the case for *h*TPH1.

In Figure 36, the effect on different pH values on melting temperature is shown. It is seen that both *ch*TPH1 and *h*TPH1 are stable in the entire pH range tested, and that pH does not have a great affect on the stability. For both proteins, pH 5.0 and 10.0 was also tested, but due to protein precipitation, the melting temperature could not be determined. The pH optimum is from 7.0 to 8.0 for both proteins, and the pH working range is from 6.0-9.0.

By comparing the maximum melting temperature of *ch*TPH1 (46 °C) and *h*TPH1 (54 °C), it is seen that the tetrameric structure or the presence of the regulatory domain in *h*TPH1 significantly stabilizes the protein.

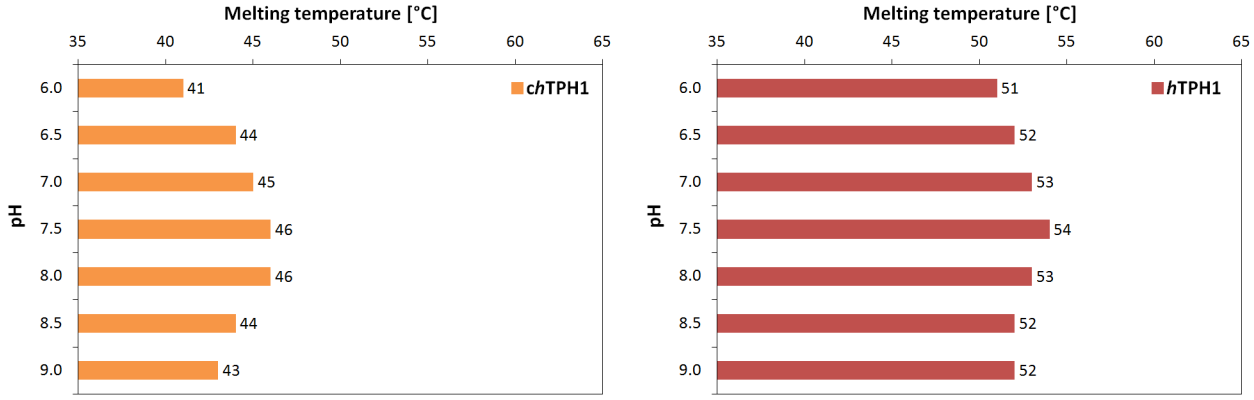


Figure 36: Results from DSF studies of the pH optimum of *ch*TPH1 (left in orange) and hTPH1 (right in red). 22 mM citric acid, 33 mM HEPES, 44 mM 2-(Cyclohexylamino)ethanesulfonic acid (CHES)/NaOH, 300 mM $(NH_4)_2SO_4$ was used as background buffer.

The effect of different salts on the stability was also tested, as seen in Figure 37. All salt concentrations were adjusted so the ionic strength matches that of 300 mM $(NH_4)_2SO_4$, meaning that the concentrations of NaCl, KCl and NH_4Cl were 900 mM.

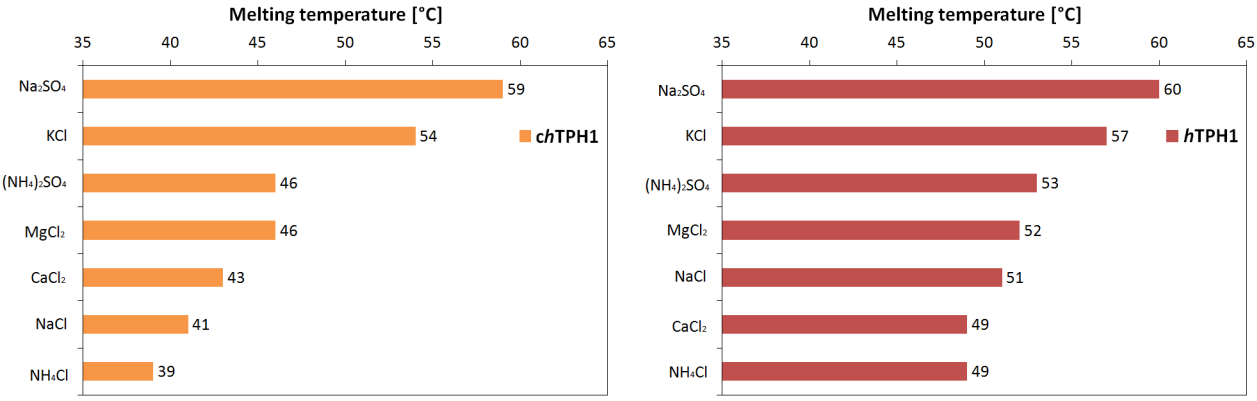


Figure 37: Results from DSF studies on the effect of different types of salt on *ch*TPH1 (left in orange) and *h*TPH1 (right in red). 100 mM Tris/H_2SO_4 or HCl, pH 8.0 was used as background buffer. All salt concentrations were adjusted so the ionic strength matched that of 300 mM $(NH_4)_2SO_4$.

For c*h*TPH1, a large effect of salt type is seen, with a difference of 20 °C in melting temperature of the most (Na_2SO_4) and least (NH_4Cl) stabilizing salt. It can be speculated that certain ions could stabilize the protein by specific interactions, but there does not seem to be any trend in the results indicating such an effect.
Based on the results above, *h*TPH1 was purified using 300 mM Na_2SO_4 instead of 300 mM $(NH_4)_2SO_4$. However, this decreased the yield of the purification from 16 to 8 mg/L culture, suggesting that the increased stabilization of *h*TPH1 does not lead to a greater yield.
The effect of the concentration of $(NH_4)_2SO_4$ was also investigated, and the results are seen in Figure 38 below. Both c*h*TPH1 and *h*TPH1 display higher melting temperatures as the $(NH_4)_2SO_4$ concentration increases. However, when studying the data closely, it is also revealed that the start fluorescence increases and the difference between the upper and lower levels of fluorescence decreases, indicating that a significant part of *h*TPH1 is precipitated or aggregated at high $(NH_4)_2SO_4$ concentrations. A purification of *h*TPH1 using 900 mM $(NH_4)_2SO_4$ correspondingly decreased the yield from 16 to 7 mg/L culture.

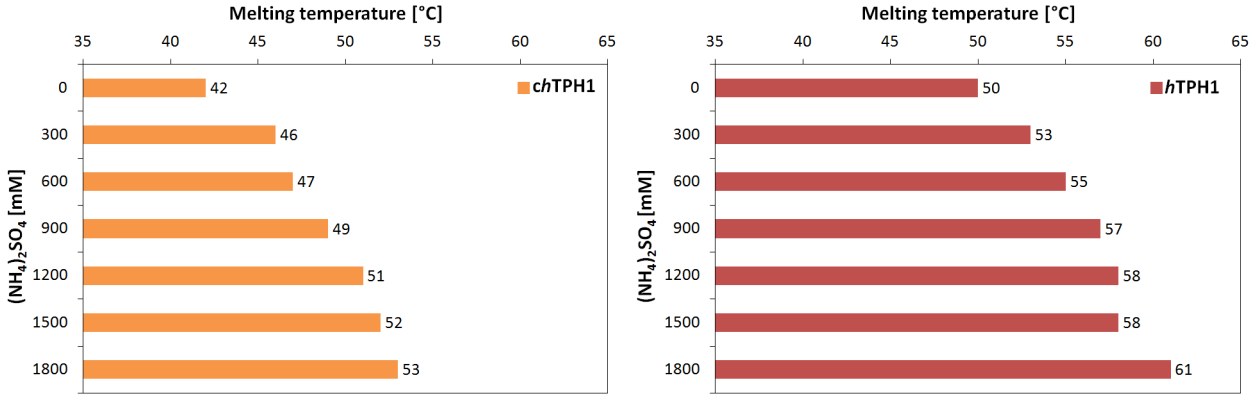


Figure 38: Results from DSF studies on the effect of different $(NH_4)_2SO_4$ concentrations on c*h*TPH1 (left in orange) and *h*TPH1 (right in red). 100 mM Tris/H_2SO_4, pH 8.0 was used as background buffer.

4.3.2.4 Purification of MBP-tagged rchTPH1

The successful purification of MBP-tagged full-length *h*TPH1 inspired to also test the purification of other variants with MBP-tag. MBP-rc*h*TPH1 was purified as described in section 4.2.3.3 generally following the same procedure as for MBP-*h*TPH1. The chromatogram from the first affinity step is not shown, while the chromatogram from the first gel-filtration step is seen in Figure 39.
As with MBP-*h*TPH1, a peak (P1) appears in the dead volume of the column corresponding to species with a size that exceeds the exclusion limit of the column. The peak contains MBP-rc*h*TPH1 and so do both P2 and P3. Due to the lack of tetramerization domain, rc*h*TPH1 is expected to be monomeric. However, comparing elution volumes of P2 (146 mL) and P3 (154 mL) with the column calibration curve, these would correspond to species with sizes of approx. 175 and 230 kDa. MBP-rc*h*TPH1 has a molecular weight of 90.6 kDa, and a monomer would be expected to elute around 175 mL. Based on this information, it seems that the regulatory domain causes some kind of oligomerization, possibly dimerization. Another possibility is that the regulatory domain is not compactly folded like globular proteins (which is what the calibration curve is based on), but maybe extends far out. This would also cause an increase in hydrodynamic radius.

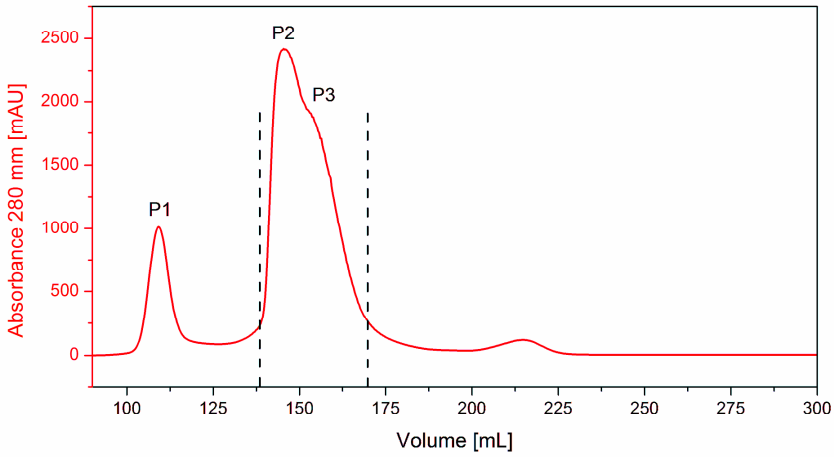


Figure 39: Chromatogram from the gel-filtration of MBP-rc*h*TPH1 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines were collected for cleavage and further purification.

Both P2 and P3 were collected and cleaved with MBP-3CP overnight. Again, the chromatogram from the second affinity step is not shown here, while the chromatogram from the second gel-filtration is shown in Figure 40. Three main peaks are seen in the chromatogram, P4, P5, and P6. P4 contains rc*h*TPH1 and elutes at 156 mL corresponding to a molecular weight of 165 kDa. P4 is the only peak in the chromatogram that displays activity in the tryptophan hydroxylation assay (data not shown). The molecular weight of rc*h*TPH1 is 47.8 kDa, and it is thus likely that rc*h*TPH1 also after cleavage exists as a dimeric or trimeric species. The P4 peak has some tailing, which could be caused by the existence of smaller species, as e.g. monomers.

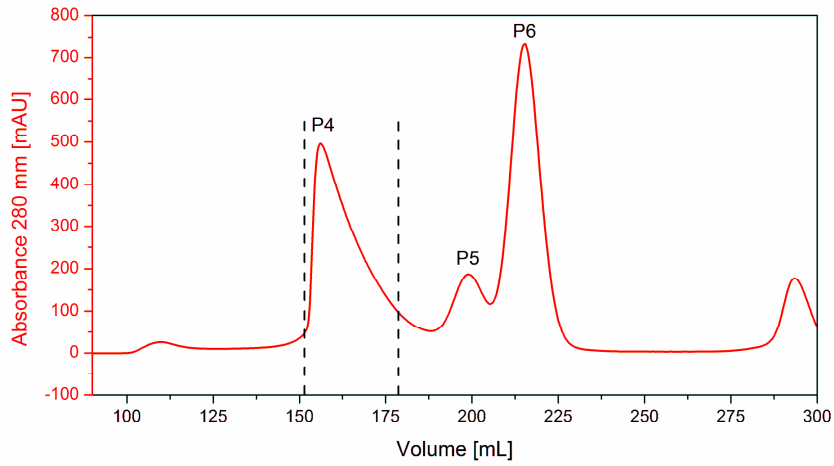


Figure 40: Chromatogram from the gel-filtration of cleaved MBP-rc*h*TPH1 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines were collected as the final product.

SDS-PAGE analysis of the purification is shown in Figure 41. It is seen that most MBP-rc*h*TPH1 is soluble and is found in the supernantant, while a small part is insoluble found in the pellet. Like for MBP-*h*TPH1, no MBP-rc*h*TPH1 is seen in the first part of the flow-through when loading on the affinity column. There is however some MBP-rc*h*TPH1 in the last part of the flow-through from the loading, which is lost. All 3 peaks (P1, P2 and P3) in the first gel-filtration chromatogram contain MBP-rc*h*TPH1, and must thus represent different oligomers of the protein. MBP-rc*h*TPH1 is completely cleaved using MBP-3CP. The last part of the

purification is somewhat difficult to follow by SDS-PAGE, as MBP has a mass of 43 kDa and rc*h*TPH1 48 kDa. They can thus be hard to separate on the SDS-PAGE.

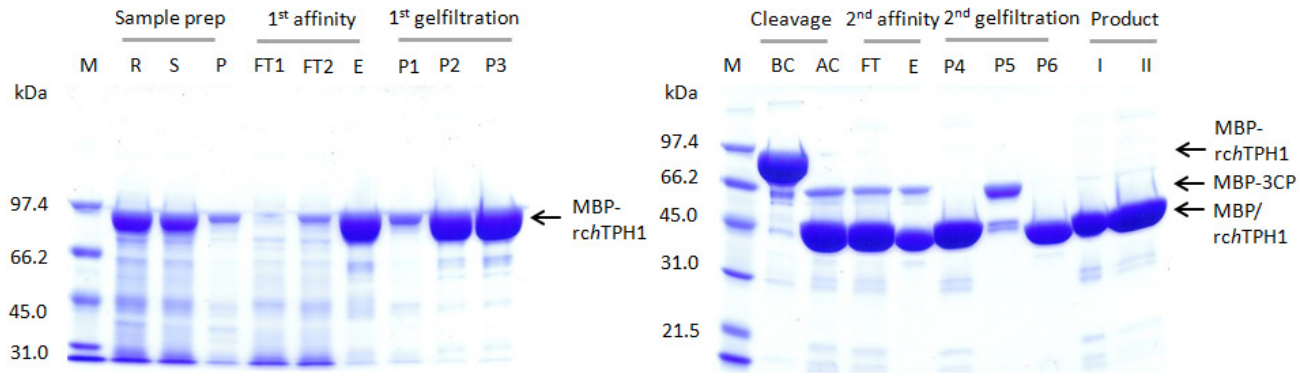


Figure 41: SDS-PAGE analysis of the purification of MBP-tagged rc*h*TPH1. Left: Purification of MBP-rc*h*TPH1. M: Molecular weight standard. R: Raw extract. S: Supernatant. P: Resuspended pellet. FT1: First part of the flow-through from the affinity step. FT2: Last part of the flow-through. E: Eluate. P1-P3: Peaks from the first gel-filtration. Right: Purification of rc*h*TPH1. BC: Before cleavage. AC: After cleavage. FT: Flow-through mixed. E: Eluate. P4-P6: Peaks from the second gel-filtration. I: The final concentrated product (1 µL). II: The final product right after gel-filtration (20 µL):

It is seen that the flow-through of the second affinity step contains impurities of MBP-3CP, which means that the step is not 100 % effective. This is also seen in the chromatogram from the gel-filtration, where several peaks appear. Only one peak (P4) showed activity. Based on this and the SDS-PAGE it is concluded that P5 contains MBP-3CP and P6 free MBP. Regarding the size and oligomerization state of rc*h*TPH1, the fact that it elutes before MBP-3CP and MBP supports the fact that rc*h*TPH1 exists as a dimer. The final product of rc*h*TPH1 is pure as judged from SDS-PAGE, but it should be noted that any potential MBP impurities would not be seen on SDS-PAGE due to the small difference in size.

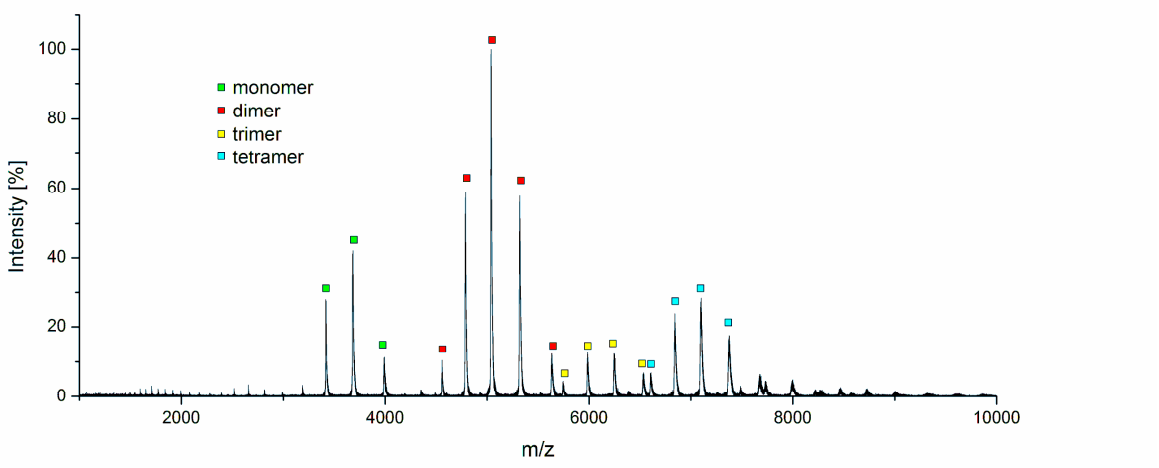


Figure 42: Mass spectrum of 23 µM rc*h*TPH1 in 1 M NH_4OAc. The calculated masses can be seen in Table 7.

Purified rc*h*TPH1 was investigated by intact mass spectrometry, in order to identify the oligomerization state of the protein. The resulting spectrum is seen in Figure 42. It is clear that rc*h*TPH1 exists in several different oligomer states in the sample. The rc*h*TPH1 dimer is the dominating species followed by monomer, tetramer and trimer. The calculated and theoretical masses of the different species can be seen in Table 7, and corresponds very well. From mass spectrometry studies the rc*h*TPH1 sample is found to be

highly polydisperse. It should be mentioned that the sample had been exchanged into 1 M NH_4OAc before mass spectrometry measurements.

Table 7: Theoretical masses of rc*h*TPH1 oligomer species including iron compared to the masses calculated from the mass spectrum of rc*h*TPH1 shown in Figure 42.

rc*h*TPH1	Theoretical mass incl. Fe [Da]	Calculated mass (from MS) [Da]
monomer	47868.3	47874
dimer	95736.6	95744
trimer	143604.9	143776
tetramer	191473.2	191699

The overall yield of the purification was in the range 16-20 mg/L culture. The yield of the purification was furthermore followed after each purification step. Half of the protein is lost in the first gel-filtration step. It seems that the yield increases drastically after the second affinity step, but this is caused by the large amount of impurities in the sample after cleavage. In the calculation of yield, all absorbance at 280 nm is assumed to be caused by rc*h*TPH1, but after cleavage both MBP and MBP-3CP will also be present. The total yield of the purification is 21 %, which is comparable to the total yield in the purification of *h*TPH1, see Table 6.

Table 8: Protein yield after each step in the purification of MBP-tagged rc*h*TPH1. The yield was calculated from the absorption at 280 nm assuming that the samples contained only MBP-rc*h*TPH1 or rc*h*TPH1 (after cleavage). The yield in % is given assuming that the first affinity step corresponds to 100 %.

	MBP-rc*h*TPH1	
	Yield	
Step	**[µmol]**	**[%]**
1st affinity step	3.8	100
1st gel-filtration step	1.9	51
2nd affinity step	3.8	99
2nd gel filtration	0.8	21

4.3.2.4.1 Monodispersity

The monodispersity of the purified rc*h*TPH1 was tested after freezing and thawing. Upon thawing, precipitation of the sample was observed, and so the precipitate was removed by filtration. The resulting chromatogram in Figure 43 shows that rc*h*TPH1 is monodisperse. This is in contrast to mass spectrometry results which showed a highly polydisperse sample. However, the mass spectrometry sample was exchanged into a different buffer for measurements, and also the vaporization process and high voltages involved in mass spectrometry could affect the oligomerization state. Based on the monodispersity test, the rc*h*TPH1 sample should be suitable for crystallization purposes. It should be noted that the high tendency to precipitate might be a problem in crystallization even though the sample is monodisperse after filtration. From the elution volume of rc*h*TPH1 and the calibration curve of the analytical column, the size should be around 80 kDa, which again indicates that rc*h*TPH1 exists as a dimer (rc*h*TPH1: 47.8 kDa).

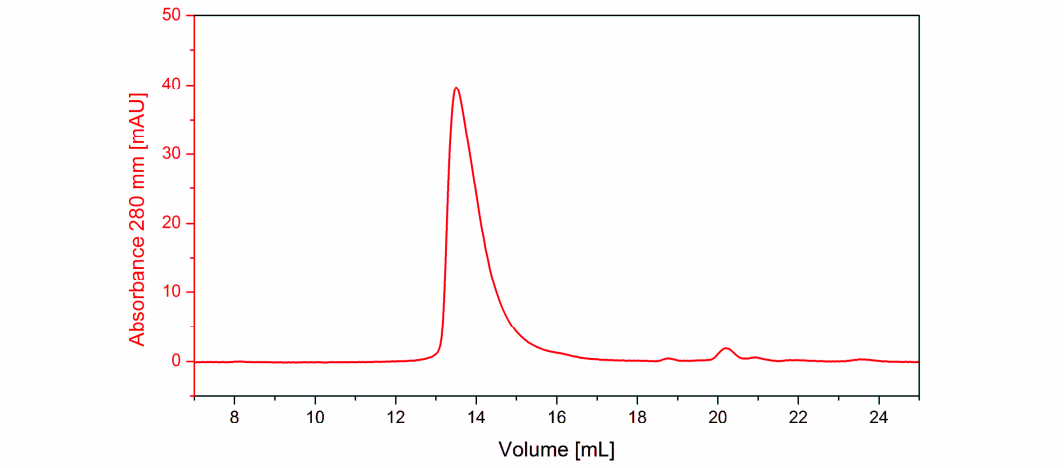


Figure 43: Chromatogram from the analytical gel-filtration of thawed rc*h*TPH1 on a superdex 200 10 300 GL column. As the sample was monodisperse, only one run was made.

4.3.3 Purification of TPH2

Variants of TPH2 were purified both with and without affinity tags. The results from the purification of c*h*TPH2 and ct*h*TPH2 were developed in earlier projects, and are only shown for future reference.

4.3.3.1 Purification of ch*TPH2*

chTPH2 was purified according to the procedure described in section 4.2.4.1 consisting of an anion-exchange step followed by a gel-filtration step. The chromatograms for both steps can be seen in Figure 44 and Figure 45, respectively. Only one major peak appears in the anion-exchange chromatogram, where c*h*TPH2 elutes at conductivity around 12 mS/cm.

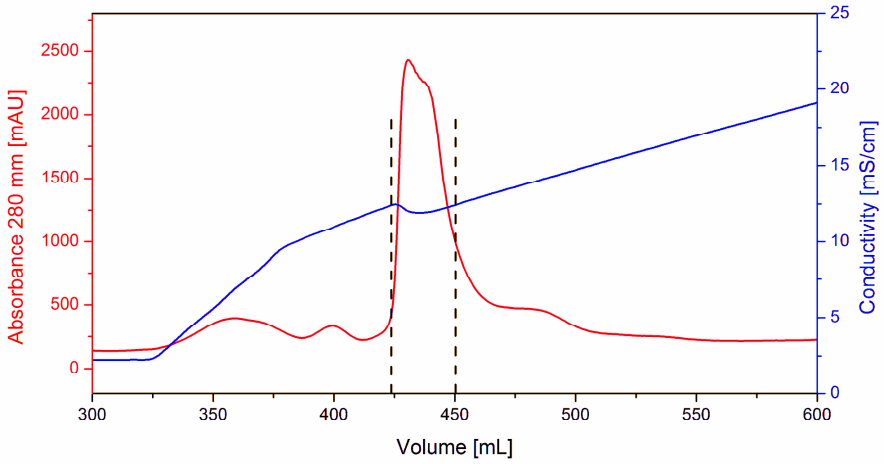


Figure 44: Chromatogram from the anion-exchange of c*h*TPH2 on a Q Seph HP 26/10 column. The fractions marked with dashed lines were collected for further purification.

The product from the anion-exchange was concentrated and loaded on a gel-filtration column. The chromatogram can be seen in Figure 45. Only one major peak appears corresponding to c*h*TPH2 monomer with a size of 36 kDa. The gel-filtration step in the purification of c*h*TPH2 has been performed in different buffers depending on the purpose that the product should be used for. The normally used buffer is 20 mM HEPES/NaOH, 100 mM $(NH_4)_2SO_4$, pH 7.2. For mass spectrometry, c*h*TPH2 has been purified in 300 mM

NH_4OAc, no pH adjustment. For crystallization, c*h*TPH2 has been purified in 20 mM HEPES/NaOH, 200 mM NaCl, pH 7.2 to avoid sulfate ions, which crystallizes with calcium ions present in many screens. The purification has also been performed with 5 (w/v) % glycerol in all buffers for an extra stabilizing effect.

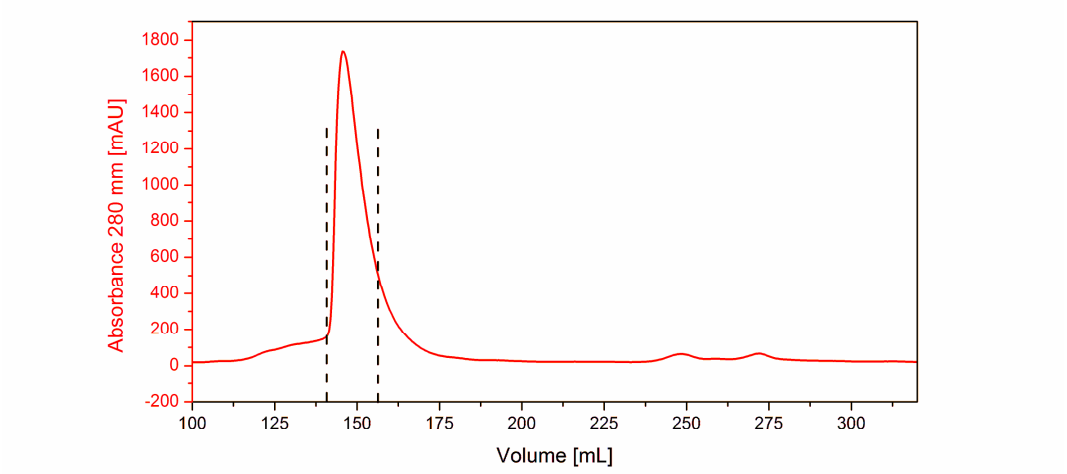


Figure 45: Chromatogram from the gel-filtration of c*h*TPH2 on a HiLoad Superdex 75 26/60 prep grade column. The fractions marked with dashed lines were collected as the final product.

Products from purifications using two different buffer systems are shown in Figure 46. The final product is pure, and thus seems suitable for crystallization. The yield of the purification is in the range 50-70 mg/L culture. The final product shows an activity of 219 ± 2.8 min^{-1}. Mass spectrometric measurements shows that the mass of the denatured protein matches exactly that of apo-c*h*TPH2 (36317 Da vs. 36316.5 Da). The mass spectrum of c*h*TPH2 can be seen in Figure 47.

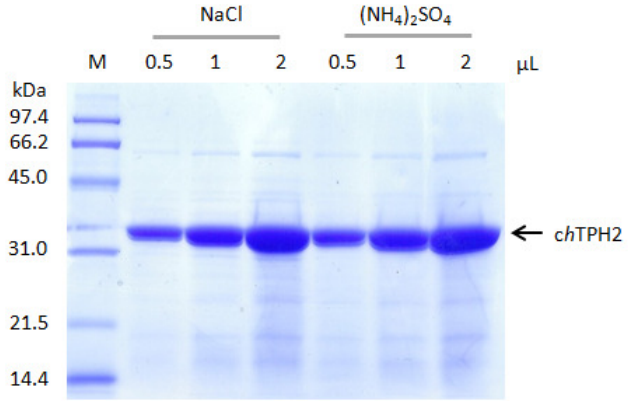


Figure 46: SDS-PAGE analysis of the final product from purifications of c*h*TPH2. M: Molecular weight standard. Numbers: µL loaded on the gel. NaCl: 20 mM HEPES/NaOH, 200 mM NaCl, pH 7.2. $(NH_4)_2SO_4$: 20 mM HEPES/NaOH, 100 mM $(NH_4)_2SO_4$, pH 7.2.

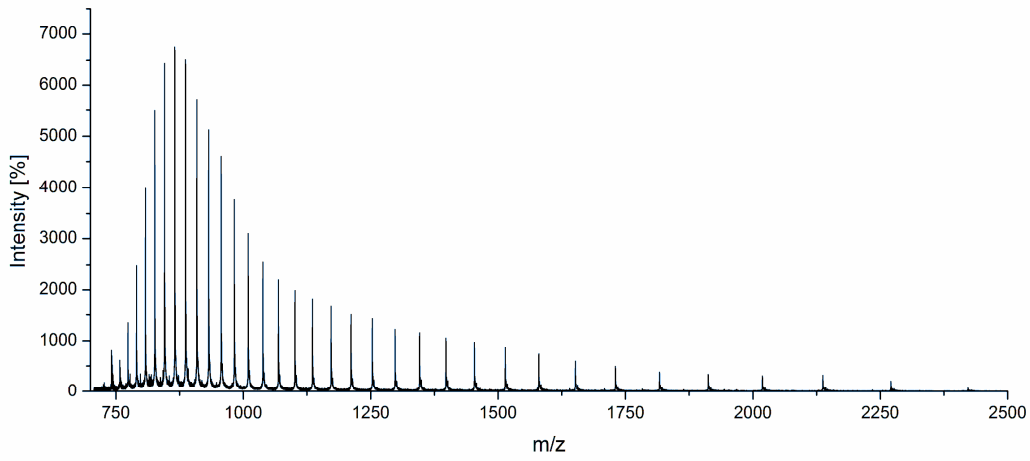


Figure 47: Mass spectrum of denatured *ch*TPH2 in 5 (v/v)% formic acid. The calculated mass is 36317 Da.

4.3.3.1.1 Monodispersity

To check the monodispersity of the final c*h*TPH2 product, it was thawed and loaded on an analytical gel-filtration column. The resulting chromatogram can be seen in Figure 48, and shows that the sample is almost monodisperse with only a small shoulder left of the main peak. The main peak was collected and loaded on the column again for a second run, which is also shown in Figure 48. Most of the small shoulder has been removed in the second run, and the sample can be considered monodisperse.

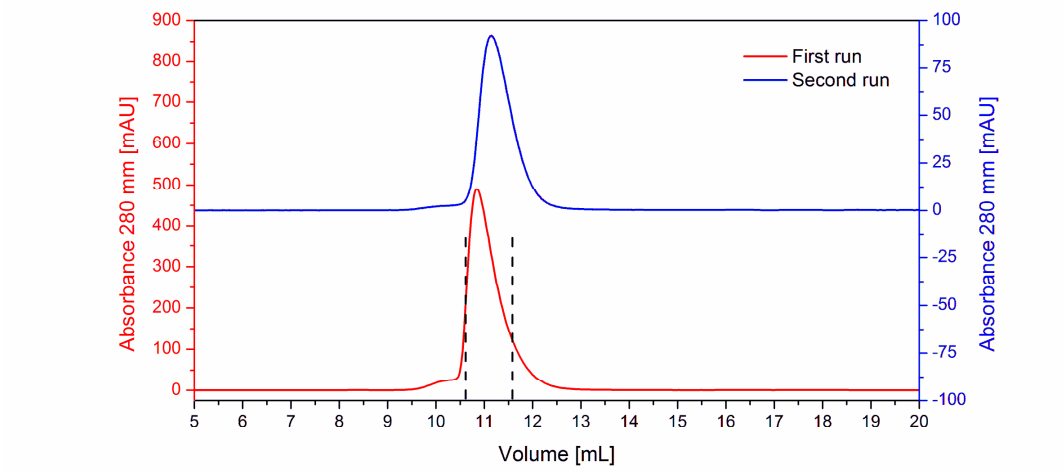


Figure 48: Chromatograms from the analytical gel-filtrations of thawed *ch*TPH2 on a superdex 75 10 300 column. Red: The first run of thawed *ch*TPH2. The fractions marked with dashed lines were collected and loaded again for the second run. Blue: The second run of *ch*TPH2 collected from the main peak in the first run. The broader peak in the second run is caused by a larger sample volume.

*4.3.3.2 Purification of ct*h*TPH2*

ct*h*TPH2 was also purified according to the same simple procedure as c*h*TPH2, with the only exceptions that a different gradient was used in the anion-exchange step and that a different gel-filtration column was used, due to the larger size of tetrameric ct*h*TPH2. In Figure 49, the chromatogram from the anion-exchange is seen. ct*h*TPH2 elutes at a conductivity around 19.5 mS/cm.

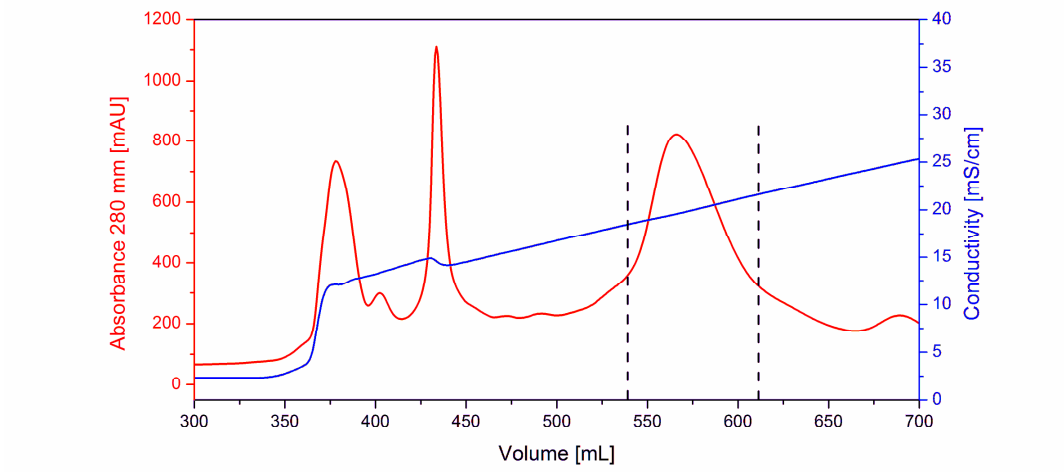


Figure 49: : Chromatogram from the anion-exchange of ct*h*TPH2 on a Q Seph HP 26/10 column. The fractions marked with dashed lines were collected for further purification.

The collected fractions were concentrated and loaded on a gel-filtration column, giving the chromatogram shown in Figure 50. The only major peak contains ct*h*TPH2 and is collected as the final product. The elution volume of ct*h*TPH2 on the gel-filtration column correspond to a mass of approx. 150 kDa, which is in agreement with the theoretical mass of a ct*h*TPH2 tetramer (158 kDa).

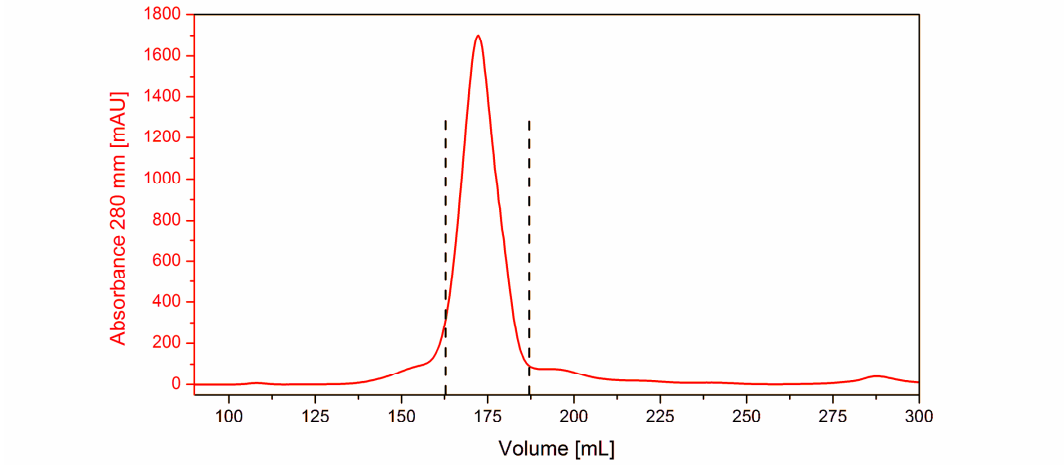


Figure 50: Chromatogram from the gel-filtration of ct*h*TPH2 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines were collected as the final product.

The last gel-filtration step was performed in several different buffers including 0.3 M NH_4OAc /NH_3, pH 7.2, 10 mM HEPES/NaOH, 200 mM NaCl, pH 7.2, and with 10 (w/v)% glycerol in all buffers for an extra stabilizing effect.

SDS-PAGE analysis of ct*h*TPH2 products from purifications using two different buffer systems are shown in Figure 51. The product is relatively pure; however some high molecular weight species are seen. These could represent impurities or ct*h*TPH2 dimers and tetramers that have not been completely dissociated in the SDS-PAGE pretreatment. The overall yield from the purification was in the range 43-54 mg/L culture, which is slightly lower than for c*h*TPH2.

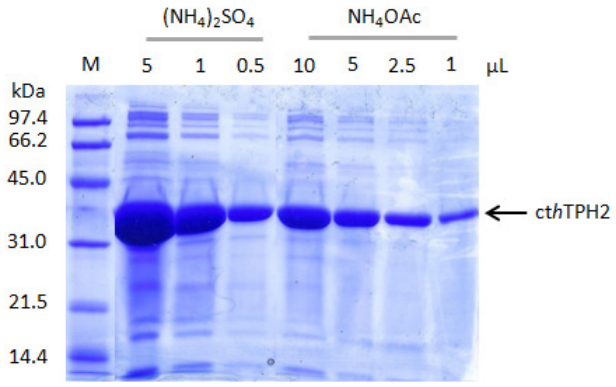


Figure 51: SDS-PAGE analysis of the final product from purifications of ct*h*TPH2. M: Molecular weight standard. Numbers: µL loaded on the gel. $(NH_4)_2SO_4$: 20 mM HEPES/NaOH, 100 mM $(NH_4)_2SO_4$, pH 7.2. NH_4OAc: 0.3 M NH_4OAc /NH_3, pH 7.2.

The final product of ct*h*TPH2 was studied by intact mass spectrometry. The resulting mass spectrum showed only one species suggesting that the sample is pure. The mass was calculated to 159187.5 Da, which corresponds well with the theoretical mass of tetrameric ct*h*TPH2 (159500.4 Da). It can thus be concluded that ct*h*TPH2 exists as a tetramer.

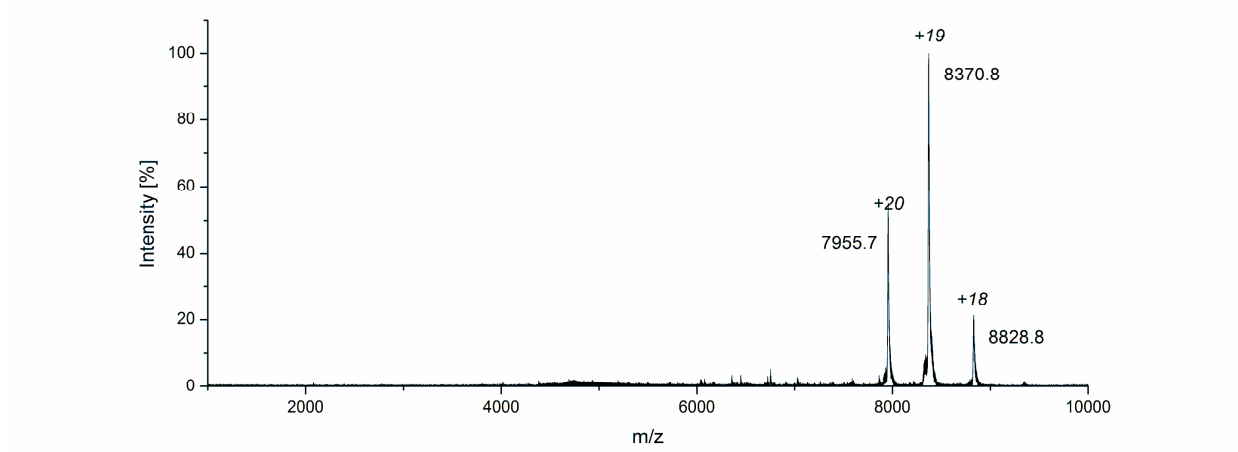


Figure 52: Mass spectrum of ct*h*TPH2 in 0.65 M NH_4OAc. The calculated mass is 159187.5 Da.

4.3.3.2.1 Monodispersity

The monodispersity of purified ct*h*TPH2 has been tested before(*11*), and it was concluded that the ct*h*TPH2 sample containing 10 % glycerol was completely monodisperse as judged by dynamic light scattering and static light scattering experiments. Static light scattering includes an analytical gel-filtration step, which showed one main peak with no shoulders. In this project, the monodispersity of purified ct*h*TPH2 with no glycerol was tested using analytical gel-filtration. The resulting chromatogram is seen in Figure 53, and shows a shoulder on the left side of the main peak. The sample without glycerol is thus not completely monodisperse.

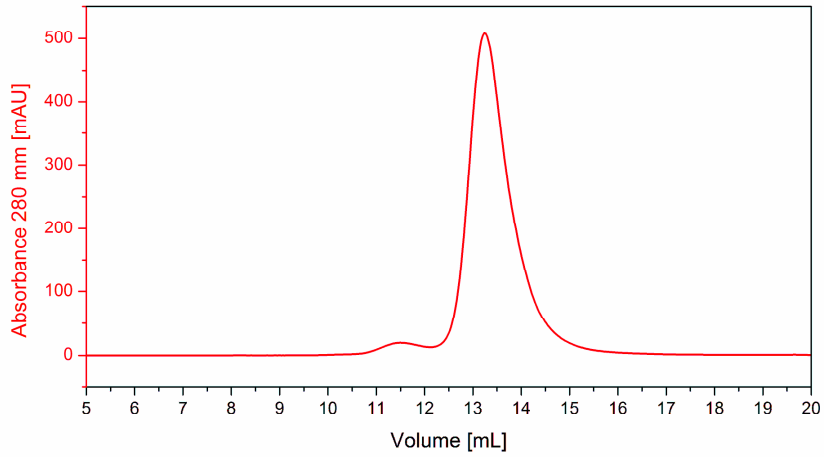


Figure 53: Chromatogram from the analytical gel-filtration of thawed ct*h*TPH2 on a superdex 200 10 300 GL column without glycerol in the buffer.

To test if presence of glycerol has an effect on monodispersity, a purification was performed with 10 (w/v)% glycerol in all buffers. The final product was frozen as usual and thawed for the monodispersity test. The result is seen in Figure 54. The shoulder is also present in the sample with 10 (w/v)% glycerol and thus glycerol does not seem to affect the monodispersity of the sample. The product from the main peak of the first run was collected and loaded again for the second run. The small shoulder on the left appeared again, suggesting that it exists in some kind of equilibrium with tetrameric ct*h*TPH2 and cannot be removed. The shoulder is however very small, so the sample can be considered almost monodisperse.

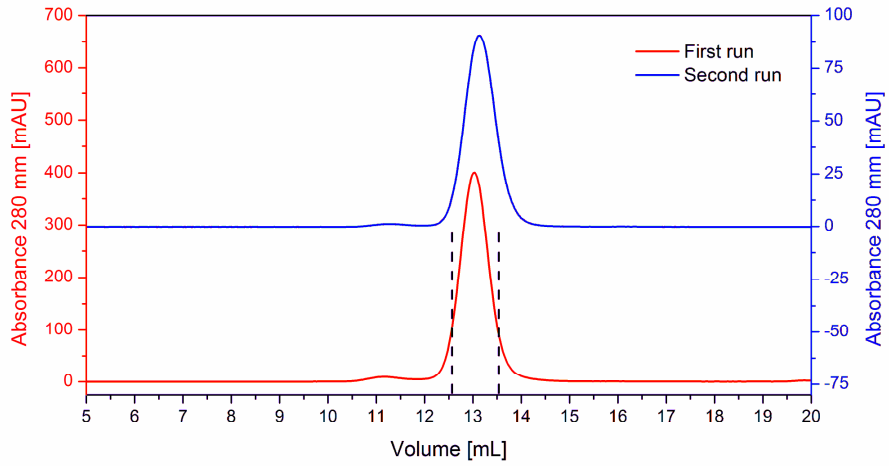


Figure 54: Chromatograms from the analytical gel-filtration of thawed ct*h*TPH2 on a superdex 200 10 300 GL column with 10 (w/v)% glycerol in the buffer. Red: The first run of thawed ct*h*TPH2. The fractions marked with dashed lines were collected and loaded again for the second run. Blue: The second run of ct*h*TPH2 collected from the main peak in the first run.

The reasons for the difference between present and past results on the monodispersity of ct*h*TPH2 are unknown. In earlier purification methods ct*h*TPH2 was put directly in the -80 °C freezer without being flash-frozen in liquid nitrogen first. The different freezing methods might affect the monodispersity.

4.3.3.3 Stability of chTPH2 and ctchTPH2

Even though successful purification protocols for c*h*TPH2 and ct*h*TPH2 have been developed, the buffers were chosen rather arbitrarily based on knowledge of theoretical pI values and some experimental evidence. By using differential scanning fluorimetry, the stability of the two TPH variants has been investigated in different buffers with varying pH, salt type and concentration. This gives a clear picture of the optimal buffer conditions, and can be used to improve the purification methods.

The melting/unfolding temperature was tested in a pH range from 5.0-9.0 using a citric acid/HEPES/CHES buffer combination with 300 mM $(NH_4)_2SO_4$. The results can be seen in Figure 55. Both c*h*TPH2 and ct*h*TPH2 have the highest melting temperature around pH 7.0. This corresponds well with the pH of 7.2 chosen for the purification methods, and is also in agreement with pH in the cytosol of cells(*120*), where TPH is mainly located(*121*). Furthermore, ct*h*TPH2 is generally more stable than c*h*TPH2 under the conditions used in the screen, as seen from the slightly higher melting temperature. An explanation could be that ct*h*TPH2 exists as a tetramer, whereas c*h*TPH2 is a monomer, and the addition of the quaternary structure could stabilize the subunits. This effect was also seen from TPH1, were *h*TPH1 was significantly more stable than c*h*TPH1. From DSF experiments it is also seen that ct*h*TPH2 is more stable at higher pH than c*h*TPH2. It can be concluded that the optimum pH is 7 for both variants, and that the pH working range for c*h*TPH2 is 6.0-7.5 and for ct*h*TPH2 it is 6.0-8.5.

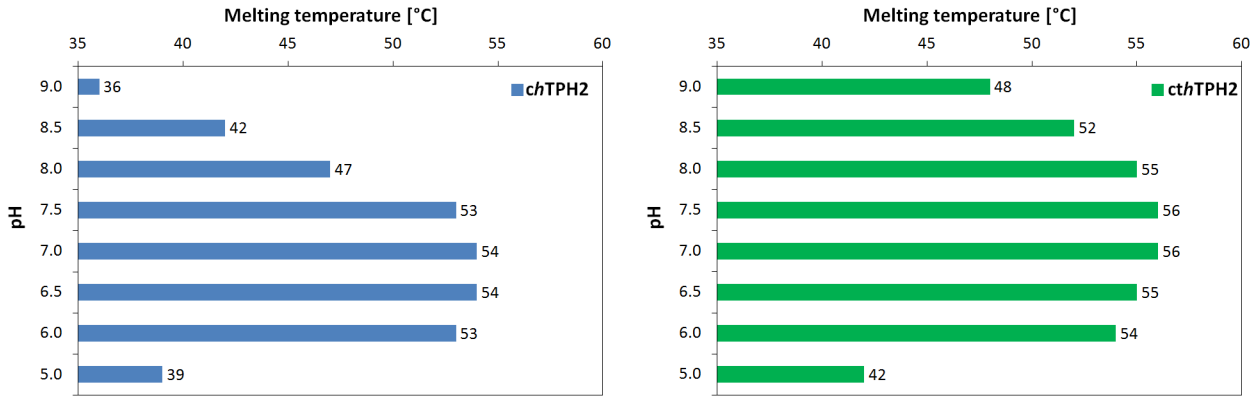


Figure 55: Results from DSF studies of the pH optimum of c*h*TPH2 (left in blue) and ct*h*TPH2 (right in green). 22 mM citric acid, 33 mM HEPES, 44 mM 2-(Cyclohexylamino)ethanesulfonic acid (CHES)/NaOH, 300 mM $(NH_4)_2SO_4$ was used as background buffer.

When comparing the stability of TPH2 variants to TPH1 variants in section 4.3.2.3, the most striking difference is the low melting temperature of c*h*TPH1 (46 °C at pH 7.5) compared to c*h*TPH2 (53 °C at pH 7.5) at high salt concentration (300 mM $(NH_4)_2SO_4$). The lower stability of c*h*TPH1 also corresponds well with the fact that c*h*TPH2 can be concentrated to 2.5 mM, while c*h*TPH1 only can be concentrated to 0.2 mM. The large difference between the catalytic domains of the two isoforms is unexpected, as they only differ at very few amino acids. The melting temperature of full-length *h*TPH1 and ct*h*TPH2 is in the same range (54 and 56 °C at pH 7.5, respectively). However, it seems that TPH2 is much more sensitive to changes in pH than TPH1, as the highest difference in melting temperature in the pH range tested is 18 °C and 14 °C for c*h*TPH2 and ct*h*TPH2, respectively, and only 5 °C and 3 °C for c*h*TPH1 and *h*TPH1, respectively.

To study the effect of different ions on the stability of c*h*TPH2 and ct*h*TPH2, several different salts were tested. The ionic strength is the same for all salts tested and corresponds to 100 mM $(NH_4)_2SO_4$.

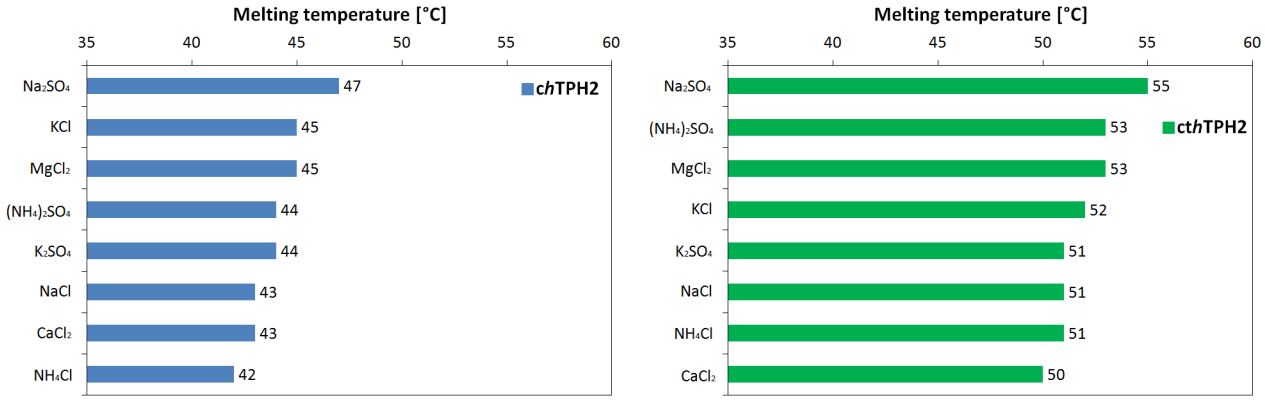


Figure 56: Results from DSF studies on the effect of different types of salt on c*h*TPH2 (left in blue) and ct*h*TPH2 (right in green). 20 mM HEPES/NaOH, pH 7.0 was used as background buffer. All salt concentrations were adjusted so the ionic strength matched that of 100 mM $(NH_4)_2SO_4$.

From Figure 56 it is seen that even though some salts seems to stabilize slightly better than $(NH_4)_2SO_4$, the melting temperature is in the same range for all salts. Thus it can be concluded that the type of ions does not have any significant effect on the stability of c*h*TPH2 and ct*h*TPH2. Whether the ionic strength has an effect was tested with 5 different $(NH_4)_2SO_4$ concentrations.

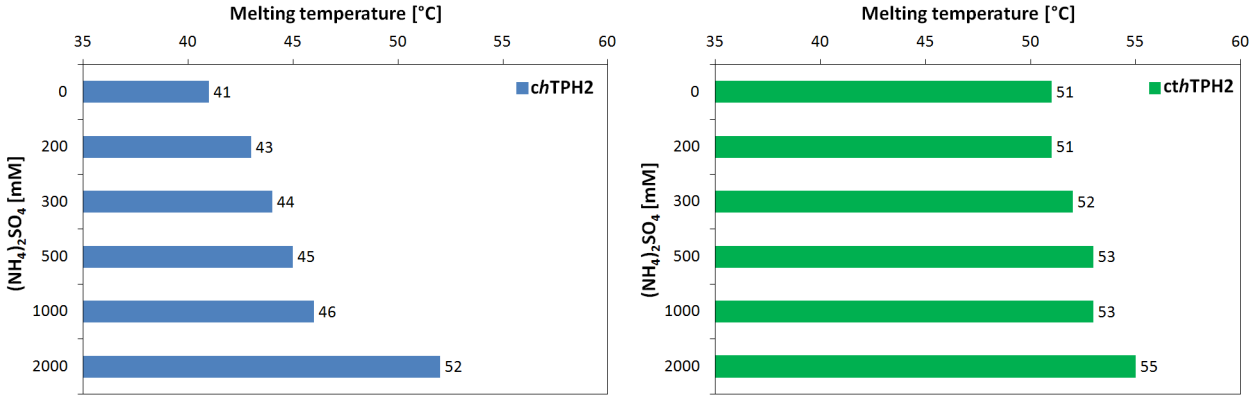


Figure 57: Results from DSF studies on the effect of different $(NH_4)_2SO_4$ concentrations on c*h*TPH2 (left in blue) and ct*h*TPH2 (right in green). 20 mM HEPES/NaOH, pH 7.0 was used as background buffer.

It is clear that the higher the ionic strength, the higher the melting temperature. However, both c*h*TPH2 and ct*h*TPH2 are stable in solution also without any salt present. It should be noted that at $(NH_4)_2SO_4$ concentrations of 300 mM the melting temperature is only 44 °C for c*h*TPH2, while the same concentration of $(NH_4)_2SO_4$ also with a pH of 7.0 but a 100 mM citric acid/HEPES/CHES combination buffer instead of just 20 mM HEPES gave 54 °C. Thus, the buffer system or concentration must have some effect on the melting temperature independent of pH and salt concentration. Compared, the melting temperature of ct*h*TPH2 is only 4 °C higher in the citric acid/HEPES/CHES combination buffer than in the HEPES only buffer (56 vs. 52 °C). To investigate this effect further three different buffers (HEPES, Bis Tris Propane and 3-(N-morpholino)propanesulfonic acid (MOPS)) with the same pH and salt concentration were tested, but showed no significant effect on stability as the melting temperature varied with only 1 °C (data not shown). Possibly, the higher buffer concentration in the combination buffer (100 mM vs. normally 20 mM) could have affected the stability.

Finally, some typical protein stabilizing molecules (glycerol, sucrose, polyethylene glycol (PEG) 400 and sorbitol) were studied and gave an increased melting temperature as expected. These stabilizers can be used as additives for future experiments.

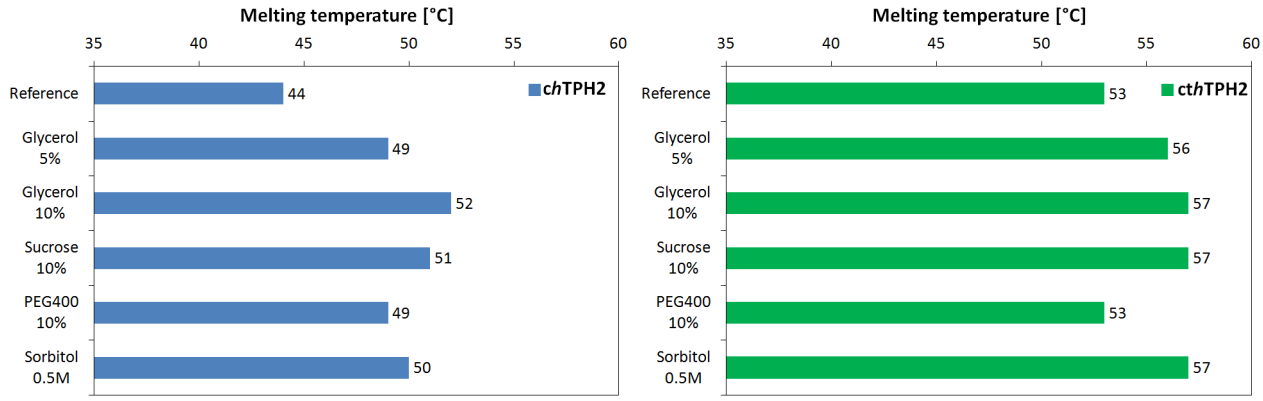


Figure 58: Results from DSF studies on the effect of different stabilizers on *ch*TPH2 (left in blue) and ct*h*TPH2 (right in green). 20 mM HEPES/NaOH, 100 mM $(NH_4)_2SO_4$, pH 7.0 was used as background buffer and reference.

4.3.3.4 Purification of h*TPH2*

In my master project, *h*TPH2 was purified using the method developed for c*h*TPH2 and ct*h*TPH2 consisting of an anion-exchange step followed by a gel-filtration step. The result was an apparently pure (judged by SDS-PAGE) and active product, however in a very low yield (4.8 mg/L culture)(*11*). Dynamic light scattering measurements of the purification product revealed several different sized species in the sample upon concentration(*11*). Thus the *h*TPH2 purification product was investigated further.

First, the chromatogram from the anion-exchange of *h*TPH2 is shown in Figure 59. *h*TPH2 elutes together with several other proteins in the first peak (see SDS-PAGE in Figure 62).

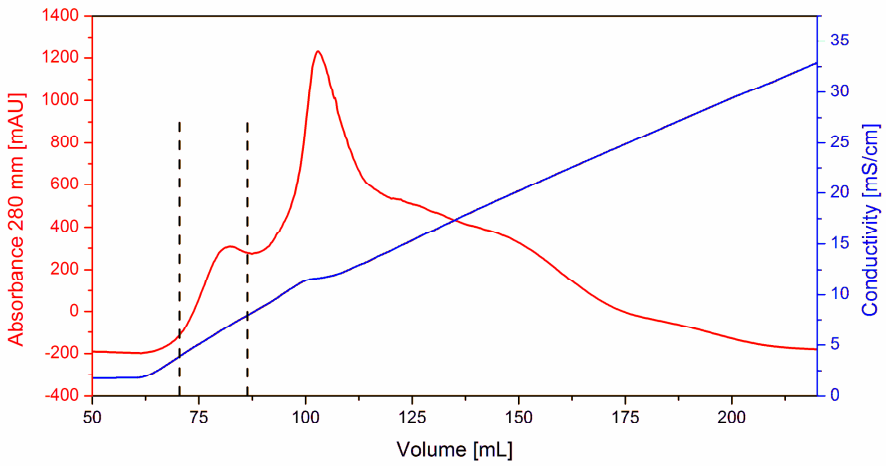


Figure 59: Chromatogram from the anion-exchange of *h*TPH2 on a Q Sepharose FF 16/10 column. The fractions marked with dashed lines were collected for further purification.

The selected fractions from the anion-exchange were concentrated and loaded onto a gel-filtration column, which resulted in the chromatogram shown in Figure 60.

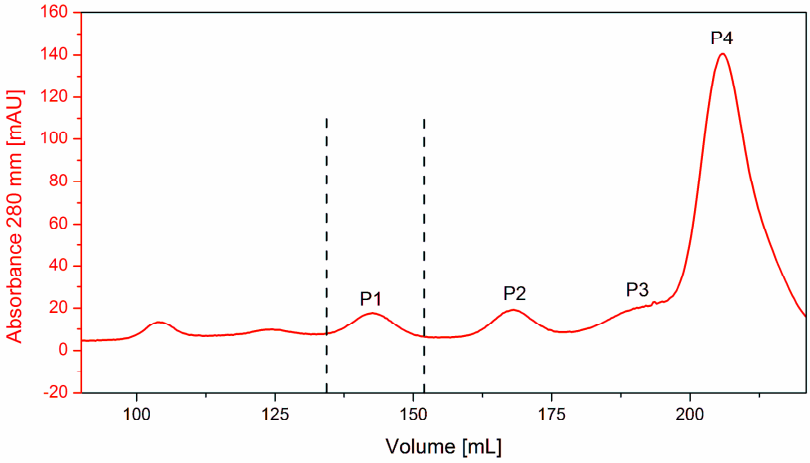


Figure 60: Chromatogram from the first gel-filtration of *h*TPH2 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines were collected and loaded on the gel-filtration column again, see Figure 61.

P1 from Figure 60 seem to contain pure *h*TPH2 (56 kDa) as judged from the SDS-PAGE shown in Figure 62. However, when the sample from P1 was concentrated and loaded on the gel-filtration column again, a chromatogram very similar to that from the first gel-filtration appeared. The chromatogram from the second gel-filtration is shown in Figure 61. The 4 main peaks from the first gel-filtration appear again in the second gel-filtration, despite the fact that only P1 was concentrated and loaded on the column. The peak intensities relative to each other are also the same in both gel-filtrations. This indicates that *h*TPH2 exists in some sort of equilibrium with other species. Activity test of the peaks from the first gel-filtration shows that only P1 displays activity, and western blotting confirms that only P1 contains TPH(*11*). Mass spectrometric identification also confirms that P1 contains *h*TPH2.

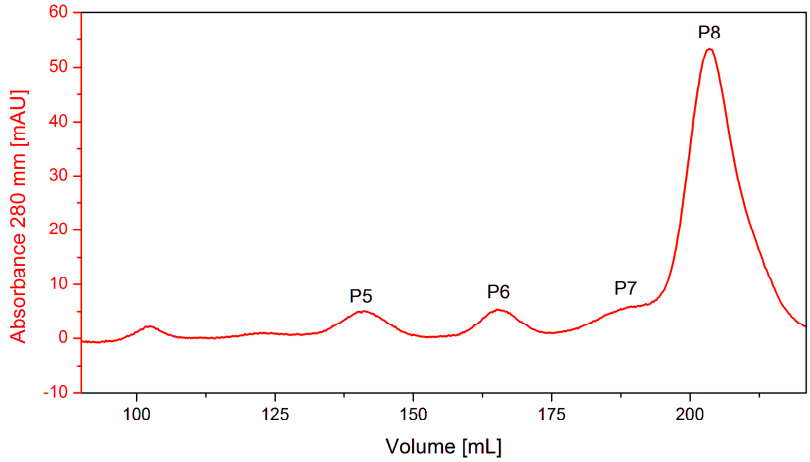


Figure 61: Chromatogram from the second gel-filtration of *h*TPH2 on a HiLoad Superdex 200 26/60 prep grade column.

SDS-PAGE analysis was used to investigate the situation further. The anion-exchange product (AP) contains at least 4 other proteins than *h*TPH2 as shown in Figure 62. The 18 % SDS-PAGE gel shows 2 proteins with sizes between 31 and 45 kDa, and two other proteins with sizes between 21.5 and 31 kDa in AP. However, P1 from the first gel-filtration appears pure and does not show any other bands than the one representing *h*TPH2 around 50 kDa. Even though the bands from the second gel-filtration are very weak, it can still be

concluded that P5 contains protein with a size of approx. 55 kDa, P6 and P7 contain approx. 40 kDa proteins, and P8 contains a protein with a size of approx. 25 kDa. The bands from the SDS-PAGE gel were cut out and send for mass spectrometric identification. The results showed that P6 contains glyceraldehyde-3-phosphate dehydrogenase (GAPD – 35.5 kDa) and P8 contains superoxide dismutase (SOD – 21.3 kDa), both from *E. coli*. GAPD catalyzes a step in glycolysis and SOD convert superoxide into oxygen and hydrogen peroxide. They are both considered "house-keeping" proteins.

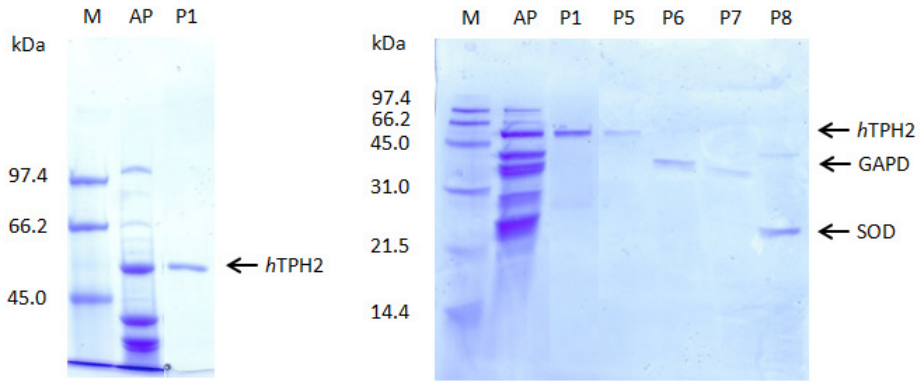


Figure 62: SDS-PAGE analysis of the purification of *h*TPH2. GAPD: glyceraldehyde-3-phosphate dehydrogenase. SOD: Superoxide dismutase. M: Molecular weight standard. AP: Anion-exchange product. P1-P8: Peaks from the first and second gel-filtration. Left: 7.5 % Tris-HCl gel. Right: 18 % Tris-HCl gel.

From the results is seems that the compound isolated from P1 in the first gel-filtration is in equilibrium with several smaller species different from TPH, which turned out to be SOD and GAPD. However, the SDS-PAGE analysis of P1 shows only one single band. The explanation could be that TPH forms a heterotetramer with a mixture of GAPD and SOD. When only one band is present on the SDS-PAGE gel, it could be because SOD and GADP linked together has almost the same mass as *h*TPH2. If SOD and GAPD are covalently linked they might not be separated in the SDS-PAGE pretreatment. According to the calibration curve of the column, the species in P1 should have a size of approx. 250 kDa corresponding with a *h*TPH2 homotetramer (224 kDa) or a mixed tetramer where each *h*TPH2 molecule could be replaced by SOD+GAPD (56.8 kDa). The two small peaks left of P1/P5 in the chromatogram also contain TPH and displays activity (data not shown). From the calibration curve, they could correspond to *h*TPH2 8-mers and 16-mers, respectively. If equilibrium exists between a mixed tetramer and the free species of SOD, GAPD and TPH in different oligomer states, it makes sense that the mixed tetramer dissociates in order to reestablish the equilibrium when isolated from the rest of the solution as in the first gel-filtration step. The exact kind of interaction between *h*TPH2, GAPD and SOD cannot be elucidated from these results. However, it appears that *h*TPH2 cannot be isolated, and thus a new purification method should be developed. Furthermore, the fact that *h*TPH2 interacts with other proteins during purification implies that it might also interact with and be stabilized by other proteins *in vivo*. This has been also been suggested in the literature(*98*), due to the instability of TPH *in vitro*.

4.3.3.5 Purification of MBP-tagged rchTPH2

As the purification of *h*TPH2 without affinity tag was unsuccessful, a new approach was taken. The success of purifying *h*TPH1 and rc*h*TPH1 with MBP as affinity tag made the attempt to purify rc*h*TPH2 and *h*TPH2 in the same way a natural choice. First, the purification of MBP-rc*h*TPH2 was studied, and the results are presented here.

In the purification of MBP-rc*h*TPH2, only one tube of cells was used (corresponding to 0.65 L culture). The purification follows the general procedure for affinity-tagged TPH variants shown in Figure 12, and the specific method described in section 4.2.4.4. The results from the DSF stability studies on c*h*TPH2 and ct*h*TPH2 described in section 4.3.3.3 shows a pH optimum around 7. Thus, pH 7 was used in all buffers in the purification of MBP-rc*h*TPH2. The MBP-column has a pH working range >7.

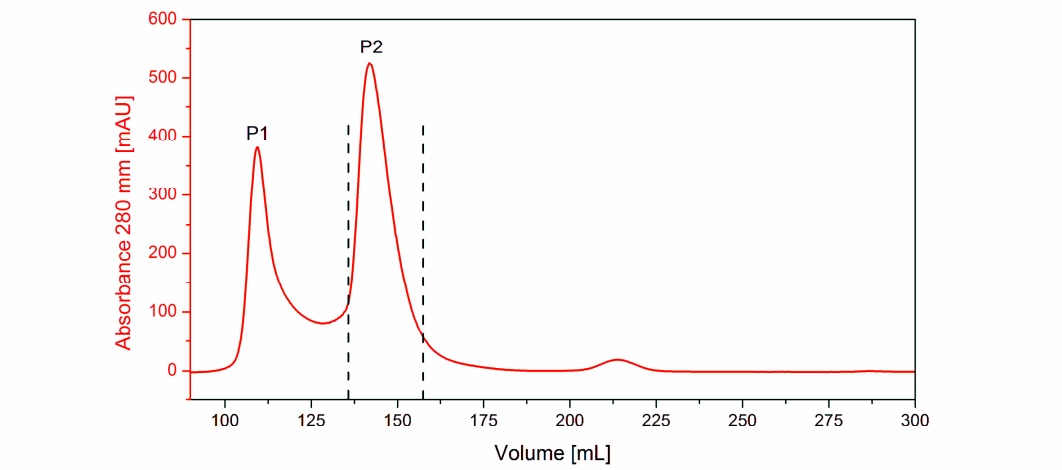


Figure 63: Chromatogram from the gel-filtration of MBP-rc*h*TPH2 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines were collected for further purification.

The first affinity step was performed on a MBPTrap HP 5 mL column. The elution fraction contained TPH and was concentrated and loaded on a gel-filtration column. The resulting chromatogram can be seen in Figure 63. As for MBP-*h*TPH1 and MBP-rc*h*TPH1, two major peaks appear in the chromatogram, one representing a high molecular weight species eluting in the dead volume of the column (P1), and one with a size corresponding to a dimer of MBP-rc*h*TPH2 (P2). P1 also contains MBP-rc*h*TPH2.

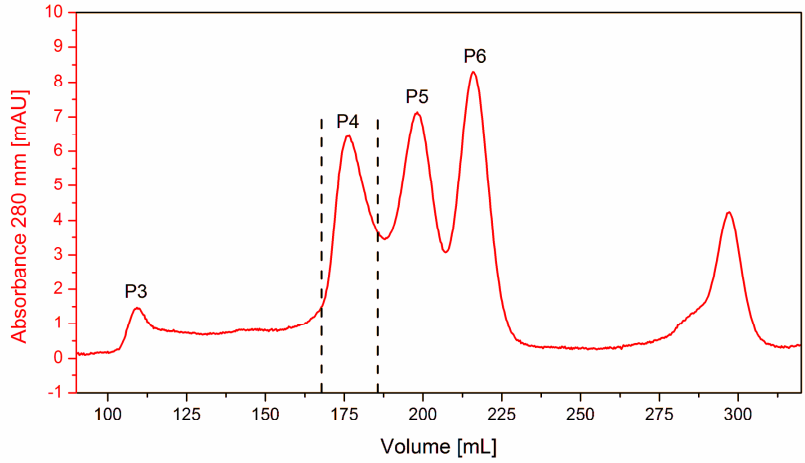


Figure 64: Chromatogram from the gel-filtration of rc*h*TPH2 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines contained mainly rc*h*TPH2.

P2 was collected and added MBP-3CP for cleavage overnight. After cleavage, precipitation was observed in the sample, so it was filtrated and then loaded on the affinity column. rc*h*TPH2 was found in the flow-

through, which was collected, concentrated and loaded on the gel-filtration column for the last purification step. The resulting chromatogram is found in Figure 64.
Several peaks appeared and from the calibration curve it was not clear which peak contained rc*h*TPH2. Thus, activity measurements were performed in order to identify the peak containing TPH.

Table 9: Results from the activity measurements of the 3 main peaks from the gel-filtration of rc*h*TPH2 shown in Figure 64. [int/min] is an arbitrary unit, which can be used for internal comparison between different samples.

Assay temperature	Blank [int/min]	P4 [int/min]	P5 [int/min]	P6 [int/min]
15 °C	5	49	23	11
26 °C	8	85	41	13

As seen in Table 9, P4 and P5 both show activity, while P6 does not. The activity of P5 is approx. half of that in P4, and could arise from the overlapping of P4 with P5. Because of the low concentration of protein, the activity measurements were performed at two temperatures to get a significant difference in activity between the inactive and active samples. From the activity measurements, it can be concluded that P4 contains rc*h*TPH2, and that P5 also contains some rc*h*TPH2, however in a lower concentration.
The purification was analyzed by SDS-PAGE as seen in Figure 65. Starting with the sample preparation, SDS-PAGE shows that approx. 70 % of MBP-rc*h*TPH2 is found in the supernatant as a soluble protein. MBP-rc*h*TPH2 is less soluble than MBP-rc*h*TPH1. In the first affinity step, only a very small amount of MBP-rc*h*TPH2 is found in the flow-through, while most binds to the column and is eluted in a pure solution. Gel-filtration separates the aggregated species of MBP-rc*h*TPH2 from the dimer. Most MBP-rc*h*TPH2 is cleaved, however a small amount is left uncleaved. During the cleavage, precipitation is seen, and the uncleaved protein might originate from this precipitate. The cleavage is rather fast and is almost complete after 10 minutes, as was also seen for cleavage of MBP-*h*TPH1. rc*h*TPH2 is found in the flow-through of the second affinity step, together with some minor impurities of MBP and MBP-3CP. The SDS-PAGE of the last gel-filtration complements the activity measurements showing that P4 indeed does contain rc*h*TPH2. P5 mainly contains MBP-3CP, but an impurity of rc*h*TPH2 can be spotted on the SDS-PAGE. This impurity must be the cause of the activity of P5. P6 contains only MBP which is in agreement with the lacking activity, see Table 9.

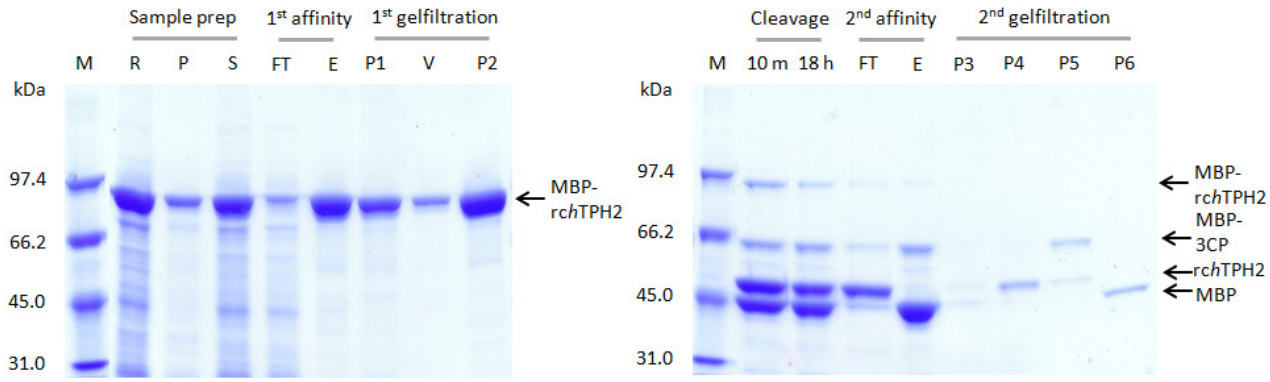


Figure 65: SDS-PAGE analysis of the purification of MBP-tagged rc*h*TPH2. M: Molecular weight standard. P: Resuspended pellet. S: Supernatant. FT: Flow-through. E: Eluate. P1-P6: Corresponding peaks in Figure 63 and Figure 64. V: Valley between P1 and P2.

It was attempted to concentrate the sample from P4 using ultrafiltration, but the protein precipitated and the final concentration was only 4.2 µM.

The purification of rc*h*TPH2 was not successful. rc*h*TPH2 cannot be separated completely from MBP-3CP in the last gel-filtration step as the resolution is too low. If a pure product should be achieved, the last affinity step should be 100 % effective removing all MBP species from the solution. Due to the instability of the final product causing precipitation, the purification was not optimized any further. c*h*TPH2 is very stable during purification and can be concentrated to 2.5 mM. The addition of the regulatory domain thus decreases the stability of the protein significantly. Furthermore, the regulatory domain apparently causes dimerization. The elution volume of rc*h*TPH2 corresponds to a molecular size of 85 kDa, and rc*h*TPH2 has a molecular weight of 53 kDa. From these results it cannot be conclusively determined if rc*h*TPH2 exists as a dimer or if the regulatory domain has a structure/fold which significantly increases the size of the protein. However, as both rc*h*TPH1 and rc*h*TPH2 apparently exist as dimers, the regulatory domains of both isoforms seem to cause dimerization. As c*h*TPH1 could be successfully purified using GST-tag, future work on purification of rc*h*TPH2 should include testing GST as affinity tag. However, the apparent dimerization of rc*h*TPH2 combined with the dimerization of GST might lead to insoluble aggregated species, as was the case for GST-*h*TPH1.

4.3.3.6 Purification of MBP-tagged h*TPH2*

As *h*TPH1 was successfully purified using MBP-tag, it was only logical to also attempt to purify *h*TPH2 using MBP-tag. The results can be seen in this section and follows the procedure described in section 4.2.4.5.

First, MBP-*h*TPH2 was loaded on the affinity column, resulting in the chromatogram shown in Figure 66. The eluate contained MBP-*h*TPH2 and was collected for further purification.

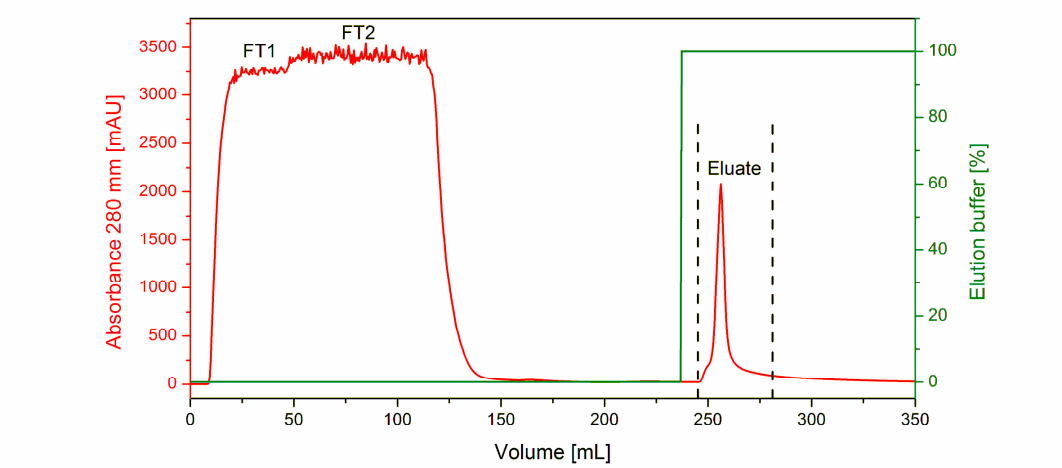


Figure 66: Chromatogram from the affinity purification of MBP-*h*TPH2 on a 20 mL Dextrin Sepharose HP 16/10 MBP column. FT1: First part of the flow-through. FT2: Last part of the flow-through. The fractions marked with dashed lines were collected for further purification.

In the flow-through an increase in absorbance is seen at approx. 45 mL, between the marks FT1 and FT2. Activity test revealed that FT1 had no activity, while FT2 showed a high activity (data not shown). This indicates that all MBP-*h*TPH2 binds in the first part of the loading process, but after 45 mL, all MBP-*h*TPH2 no longer binds and is thus found in the flow-through. SDS-PAGE confirms this hypothesis showing no bands for MBP-*h*TPH2 in FT1, while a clear band is seen in FT2 (*vide infra*, see Figure 69). To test if the MBP-

column had simply been overloaded, the FT was loaded on the column again in a new run, but nothing bound to the column and no elution peak was seen. This behavior was also seen in the purification of MBP-*h*TPH1 and all other MBP-tagged TPH variants. No explanation for the phenomenon has been found.
Despite the great loss of protein in this first step, the eluate was concentrated and loaded on a gel-filtration column. The chromatogram seen in Figure 67 shows two major peaks (P1 and P2), as for the other MBP-tagged TPH variants, and one minor peak at higher elution volume (P3). P2 which corresponds to tetrameric MBP-*h*TPH2 was collected and cleaved using MBP-3CP. The cleaved sample was loaded on the affinity column, and *h*TPH2 was collected from the flow-through.

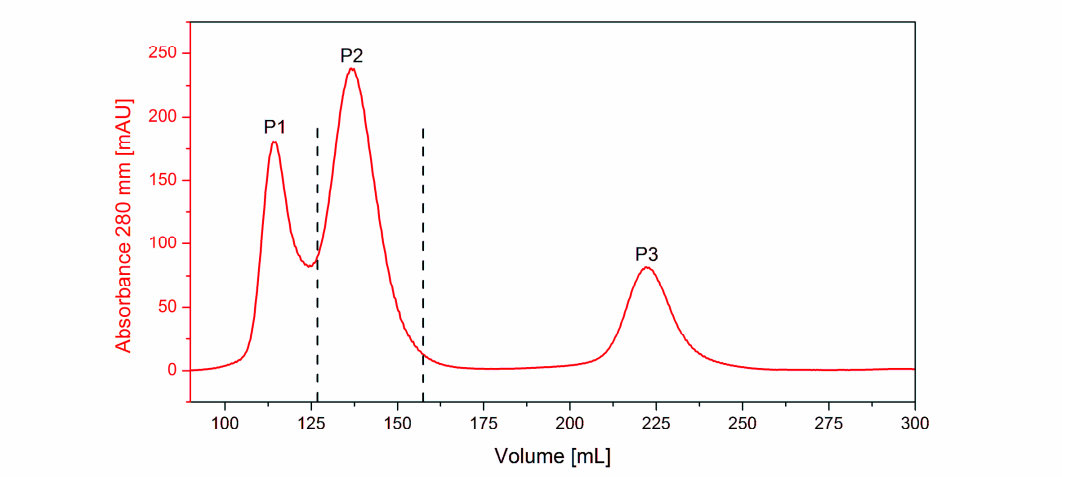


Figure 67: Chromatogram from the gel-filtration of MBP-*h*TPH2 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines were collected for further purification.

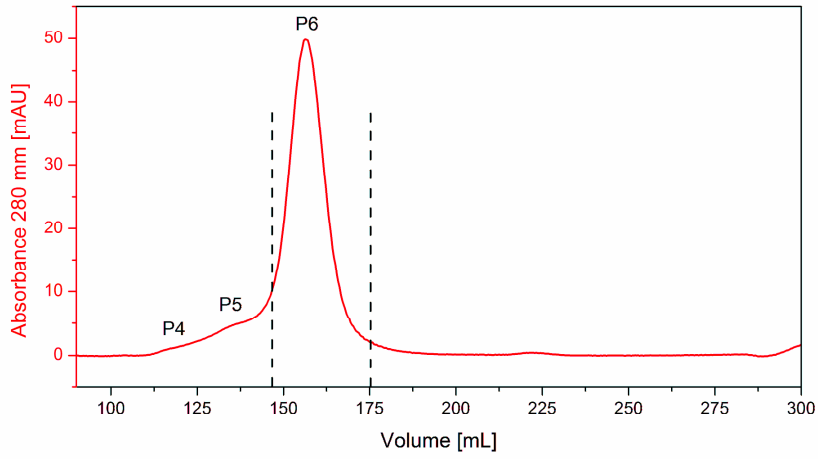


Figure 68: Chromatogram from the gel-filtration of cleaved MBP-*h*TPH2 on a HiLoad Superdex 200 26/60 prep grade column. The fractions marked with dashed lines were collected as the final product.

The flow-through was concentrated and loaded on the gel-filtration column again in the last purification step giving the chromatogram shown in Figure 68. Only one major peak is seen, and according to the elution volume it contains tetrameric *h*TPH2. The peak showed significant activity in the activity assay. The collected fractions from this peak were concentrated using ultrafiltration, but only up to 9 µM. The final yield of the purification of *h*TPH2 was 0.6-4 mg/L culture, which is a very low yield especially compared to *h*TPH1 (16 mg/L culture).

The SDS-PAGE analysis of the entire purification is seen in Figure 69. First, the solubility of MBP-*h*TPH2 is even less than for MBP-r*ch*TPH2 and significantly worse than MBP-*h*TPH1. Approx. half of the expressed MBP-*h*TPH2 is found in the supernatant, while the rest is insoluble and found in the pellet. In the first affinity step, the first part of the flow-through does not contain MBP-*h*TPH2, while the last part does, as described earlier. The reason for this lack of binding is not known. However, some MBP-*h*TPH2 is bound and eluted with minor impurities. The first gel-filtration has MBP-*h*TPH2 in both P1 and P2. The cleavage is almost complete, and *h*TPH2 elutes in the FT of the second affinity step without any visible impurities. The small shoulder, P5, left of the main peak in the last gel-filtration, contains *h*TPH2 as does the main peak P6. The latter was collected and concentrated, and from the SDS-PAGE analysis the final product of *h*TPH2 appears pure.

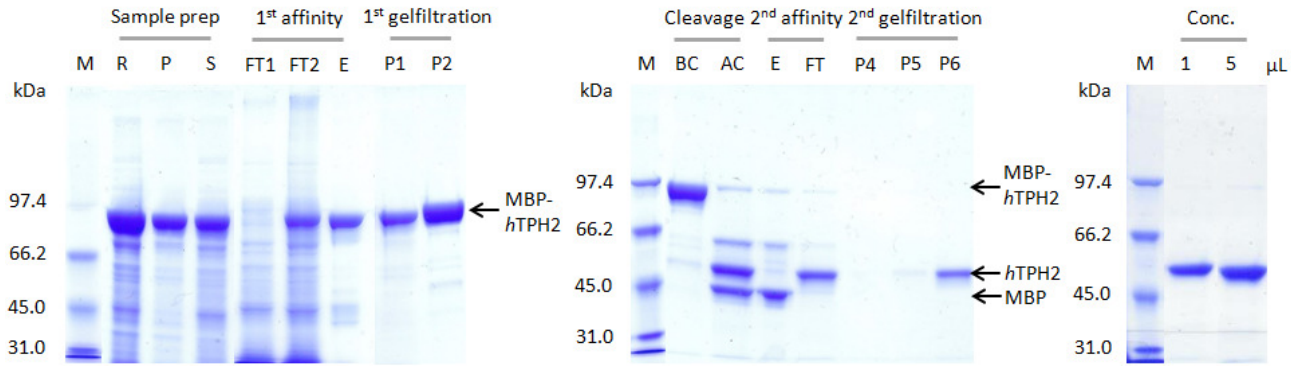


Figure 69: SDS-PAGE analysis of the purification of MBP-*h*TPH2. M: Molecular weight standard. R: Raw extract. P: Resuspended pellet. S: Supernatant. FT1 and FT2: Flow-through fractions as defined in Figure 66. E: Eluate. P1-P6: Peaks in the gel-filtration chromatograms in Figure 67 and Figure 68. BC: Before cleavage. AC: After cleavage. Numbers refers to µL loaded on the gel.

The instability of *h*TPH2 and hence the low yield gave rise to several optimization trials, but none gave positive results. The yield of each purification step for the original purification procedure can be seen in Table 10.

Table 10: The protein yield after each step in the purification of MBP-*h*TPH2. The yield was calculated from the absorption at 280 nm assuming that the samples contained only MBP-*h*TPH2 or *h*TPH2 (after cleavage). The yield in % is given assuming that the first affinity step corresponds to 100 %.

	MBP-*h*TPH2	
	Yield	
Step	**[µmol]**	**[%]**
1st affinity	0.32	100
1st gel-filtration	0.11	35
2nd affinity	0.14	44
2nd gel-filtration	0.02	6

Compared to the yields for MBP-*h*TPH1 seen in Table 6, the percentage yields are comparable. However, the amount of µmol protein is approx. a factor 10 smaller for MBP-*h*TPH2. This indicates that the first affinity step where large amount of MBP-*h*TPH2 is lost also is the main factor in limiting the yield. The yield from the first affinity step is significantly lower that the expected binding capacity of the MBP column (2.5 µmol). As the *h*TPH2 sample has no impurities after the second affinity step (shown by SDS-PAGE) the last gel-filtration step can be omitted. This would also increase the final yield.

In conclusion, the fusion to MBP allows purification of active *h*TPH2 in a pure solution. Large amounts of MBP-*h*TPH2 are lost in the flow-through of the first affinity step for unknown reasons. *h*TPH2 without affinity tag is highly unstable and precipitates easily, and cannot be concentrated to more than 9 µM. The yield of the purification was 0.6-4 mg/L culture.

4.4 Overall discussion

Several different variants of TPH have successfully been purified in this project. Here, a comparison between the variants is given in terms of expression levels, solubility, yields and activity.

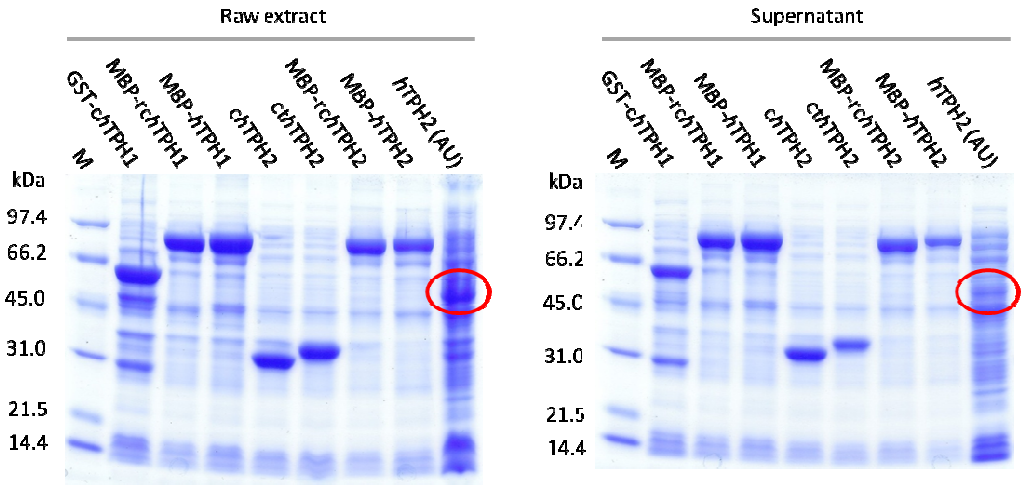


Figure 70. SDS-PAGE analysis of the raw extracts (left) and supernatants (right) of TPH variants purified in this project. *h*TPH2 was expressed in auto-inducing media (AU). The red circle marks the band representing *h*TPH2. For all other variants, the most significant band contains target protein.

The SDS-PAGE analysis of the raw extracts of all variants purified in this project in Figure 70 shows that the expression levels are comparable. It seems that TPH1 variants are expressed to a slightly higher level than TPH2 variants, independent of affinity-tags. The *h*TPH2 sample has a higher level of general protein expression, but here auto-inducing media was also used. From the supernatants and the raw extracts, the solubility of the variants can be compared. MBP-*h*TPH1, MBP-rc*h*TPH1, MBP-rc*h*TPH2 and c*h*TPH2 are all highly soluble (80-90 % of the expressed protein is soluble). ct*h*TPH2, GST-c*h*TPH1 and MBP-*h*TPH2 are partly soluble (50-70 % soluble). *h*TPH2 without affinity tag grown in auto-inducing media is only approx. 10 % soluble. For *h*TPH2, the addition of the MBP affinity tag significantly increases the soluble fraction of the protein, which was also the case for *h*TPH1, as can be seen from comparison between Figure 24 and Figure 25.

Even though significant amounts of soluble protein are available for all TPH variants, the yields of the purifications differ, as seen from Table 11. Purification of c*h*TPH2 and ct*h*TPH2 gives the highest yield in terms of mg protein per L culture. This method only involves two steps. For all TPH1 variants, the yield is somewhat lower, but still each purification procedure provides a satisfactory amount of protein, which can be concentrated to reasonable concentrations for characterization purposes. The addition of the regulatory domain does not seem to have any negative effect on the yields for TPH1 variants. However, purification of both rc*h*TPH2 and full-length *h*TPH2 gives very low yields, compared to the purifications of both c*h*TPH2 and ct*h*TPH2 and MBP-rc*h*TPH1 and MBP-*h*TPH1. The regulatory domains of TPH2 thus significantly affect the yields. Also, the TPH2 variants containing the regulatory domain has a high tendency to precipitate and

can only be concentrated to very low concentrations. It seems that the high salt concentration stabilizes the regulatory domain of TPH1, but does not give the same stabilizing effect for the regulatory domain of TPH2. The two isoforms differ significantly in the regulatory domain, and thus it is not unlikely that the two regulatory domains are stabilized by different factors.

Table 11: Overview of the purification method, yield and final concentration for the TPH variants purified in this project. *: The rc*h*TPH2 sample was not completely pure. **: Activity was measured on a different day using different protein concentration compared to the other variants.

Variant	Purification method	Yield [mg/L culture]	Concentration [mM]	Activity [min^{-1}]
c*h*TPH1	Affinity using GST-tag and gel-filtration	11	0.18 mM	92 ± 2.2
***h*TPH1**	Affinity using MBP-tag and gel-filtration	23	0.16 mM	164 ± 3.1
rc*h*TPH1	Affinity using MBP-tag and gel-filtration	20	0.20 mM	95 ± 4.2
c*h*TPH2	Anion-exchange and gel filtration	70	2.5 mM	219 ± 2.8
ct*h*TPH2	Anion-exchange and gel-filtration	54	0.29 mM	187 ± 7.5
rc*h*TPH2*	Affinity using MBP-tag and gel-filtration	0.06	0.004 mM	16**
***h*TPH2**	Affinity using MBP-tag and gel-filtration	0.6	0.009 mM	29 ± 8.3**

The tryptophan hydroxylation activities of the purified variants were also measured as described in section 4.2.5.2, and the results can be seen in Table 11 and Figure 71. The activities of rc*h*TPH2 and *h*TPH2 were measured on different days with different protein concentrations, while all other variants were measured on the same day using a protein concentration of 1.5 µM. rc*h*TPH2 and *h*TPH2 displayed activity compared to blank samples, but the protein concentrations were very low, and the protein was unstable and precipitated easily, so the results for these two should not be compared to the other variants.

The most striking result from the activity measurements is the difference in specific activity between c*h*TPH1 (92 min^{-1}) and c*h*TPH2 (219 min^{-1}). The catalytic domains of TPH1 and TPH2 have a sequence identity of 81 % and a sequence homology of 91 %(*13*), and thus it is intriguing to observe such differences in specific activity. However, this observation has been made before(*13*). Another interesting feature is the lack of effect of the regulatory domain on activity of TPH1. c*h*TPH1 and rc*h*TPH1 have similar specific activities, despite the fact that rc*h*TPH1 exists as a dimer and includes the regulatory domain. This result suggests that the regulatory domain of TPH1 does not directly influence catalytic activity in its native state, but may do so upon phosphorylation or binding of regulators. The higher specific activity of *h*TPH1 compared to c*h*TPH1 and rc*h*TPH1 indicates that the tetrameric structure improves the catalytic activity of TPH1. This could be caused by e.g. cooperativity between different subunits. However, for TPH2, the addition of the tetrameric structure slightly lowers the specific activity for ct*h*TPH2 (187 min^{-1}) as compared to c*h*TPH2 (219 min^{-1}).

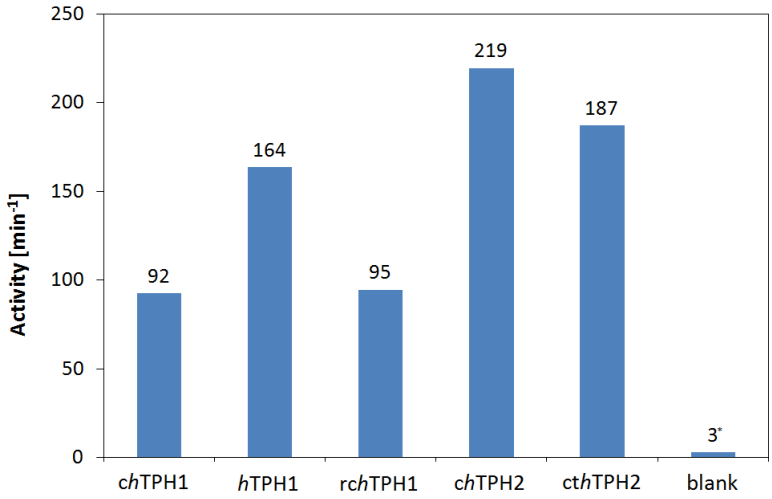


Figure 71: The specific activity of different TPH variants measured by activity assay. All measurements were made on the same day and the protein concentration was 1.5 µM for all variants. *: The unit of the blank sample is [int/min].

4.5 Conclusions

Successful purification procedures were developed for a number of TPH variants. A new purification method for c*h*TPH1 using GST affinity-tag was developed, which eliminated impurities identified by mass spectrometry in earlier projects. Furthermore, purification methods for rc*h*TPH1 and *h*TPH1 using MBP affinity-tag were developed. It is the first time a purification method which yields pure, active and stable *h*TPH1 has been presented. This allows for a thorough characterization of the full-length human enzyme, which has not been possible before. *h*TPH1 displayed a significantly higher specific activity than c*h*TPH1 and rc*h*TPH1, suggesting that the tetramerization increase reactivity, possibly by cooperativity. Furthermore, rc*h*TPH2 and *h*TPH2 were attempted purified using the same general method as for *h*TPH1, but the yields were very low and the products were unstable and could not be concentrated. The stability of c*h*TPH1, *h*TPH1, c*h*TPH2 and ct*h*TPH2 were studied by differential scanning fluorimetry, and the results gave information on the effect of pH, salt type and ionic strength.

5 Crystallization

Despite the recent development of high-throughput crystallization methods and impressive increase in available protein structures, production of high-quality crystals remains a major bottleneck in structure determination(*122*). Even when pure, soluble protein is available, the production of crystals which diffract can be a hard and sometimes impossible job, because crystallization is a complex multi-parametric process. So far, there are no complete guidelines of how to achieve well-diffracting crystals, and most of the advice found in the literature is based on empiric observations(*95*). The first step in crystallization is to scan a variety of precipitation agents, to find hits or leads, which may be optimized to produce crystals. These leads can be crystals, crystalline precipitate or even phase separation(*122*). When a lead is identified, the crystallization conditions can be fine-tuned by varying protein concentration, type and concentration of precipitant, pH or temperature or by adding additives as e.g. substrates. The most commonly used screens for initial screening of crystallization conditions are sparse matrix screens that consist of compilations of conditions that have previously led to successful crystallization(*123*). It is estimated that <300 trials for each protein should be sufficient to find crystallization conditions(*124*). Instead of testing even more conditions, the time and resources should rather be put into testing for different organisms, optimizing the constructs or looking for orthologs. When working with proteins that have potential as drug targets, it can be important to keep focus on the human variant, and instead try to optimize different constructs. With the use of high-throughput robotics, a vast amount of conditions can easily be tested without wasting large resources, time or protein, and thus this technique has also been used in this project. When a concentrated effort is put into the crystallization of a protein, the overall success rate is estimated to 30-40 %(*125*).
Even when protein crystals have been produced, there are still many obstacles before the structure can be determined. The crystals should be mounted and tested for diffraction. The collection of high-quality diffraction data is yet another bottleneck in protein crystallization, where many crystals fail, either in terms of no diffraction at all or by giving such low-quality diffraction data in terms of e.g. resolution or completeness, that the structure cannot be solved. However, when protein crystallization does succeed, the results can be of great importance for science.
This chapter will present the attempts to make a structural characterization of some of the purified variants of TPH by crystallization. First, an introduction to the already published structures of TPH will be given.

5.1 Introduction and background

Due to the difficulties in purification of TPH and problems with instability of the full-length protein, the structural characterization in form of crystal structures is very limited. Only crystal structures of the catalytic domain of isoform 1 exist. In comparison, PAH and TH has been characterized to a greater extent including structures of both the tetramerization and regulatory domains. However, the structure of a full-length AAAH has not yet been determined. In Table 12 an overview of all the published crystal structures of the AAAHs found in the PDB database can be seen.

Table 12: Overview of all structures of the aromatic amino acid hydroxylases found in the PDB database. r: Regulatory domain. c: Catalytic domain. t: Tetramerization domain. *Cv*: Chromobacterium violaceum. *Cp*: Colwellia psychrerythraea 34H.

PDB ID	Organism	Domains	Resolution	Comments	Ref.
Tryptophan hydroxylase 1					
1MLW	*Human*	c 102-402	1.7 Å	Fe(III), BH_2	(*47*)
3E2T	*Chicken*	c 101-414	1.9 Å	Fe(III), trp, imidazole	(*12*)
3HF8	*Human*	c 104-393	1.9 Å	Fe(III), inhibitor LP-533401	(*48*)
3HFB	*Human*	c 104-393	1.9 Å	Fe(III), inhibitor LP-534193	(*48*)
3HF6	*Human*	c 104-393	1.8 Å	Fe(III), inhibitor LP-521834	(*126*)
Tyrosine hydroxylase					
1TOH	*Rat*	ct 156-177, 200-498	2.3 Å	Fe(III)	(*63*)
2TOH	*Rat*	ct 160-182, 186-498	2.3 Å	Fe(III), BH_2	(*127*)
2XSN	*Human*	ct 193-527	2.7 Å	Zn(II), 8 mutations	
Phenylalanine hydroxylase					
1PAH	*Human*	c 117-424	2.0 Å	Fe(III)	(*128*)
3PAH	*Human*	c 117-424	2.0 Å	Fe(III), adrenaline	(*129*)
4PAH	*Human*	c 117-424	2.0 Å	Fe(III), noradrenaline	(*129*)
5PAH	*Human*	c 117-424	2.0 Å	Fe(III), dopamine	(*129*)
6PAH	*Human*	c 117-424	2.0 Å	Fe(III), dopa	(*129*)
1PHZ	*Rat*	rc 19-136, 143-427	2.2 Å	Fe(III), phosphorylated	(*49*)
2PHM	*Rat*	rc 19-136, 143-427	2.6 Å	Fe(III), dephosphorylated	(*49*)
2PAH	*Human*	ct 118-136, 143-452	3.1 Å	Fe(III)	(*64*)
1DMW	*Human*	c 118-424	2.0 Å	Fe(III), BH_2	(*130*)
1J8T	*Human*	c 118-424	1.7 Å	Fe(II)	(*65*)
1J8U	*Human*	c 118-424	1.5 Å	Fe(II), BH_4	(*65*)
1LRM	*Human*	c 118-424	2.1 Å	Fe(II), BH_2	
1LTU	*Cv*	2-285	1.7 Å		(*131*)
1LTV	*Cv*	9-283	2.0 Å	Fe(III)	(*131*)
1LTZ	*Cv*	7-253, 257-284	1.4 Å	Fe(III), BH_2	(*131*)
1KWO	*Human*	c 118-424	2.5 Å	Fe(II), BH_4, thienylalanine	(*132*)
1MMK	*Human*	c 117-425	2.0 Å	Fe(II), BH_4, thienylalanine	(*66*)
1MMT	*Human*	c 117-424	2.0 Å	Fe(II), BH_4, norleucine	(*66*)
1TDW	*Human*	c 117-424	2.1 Å	Fe(III), A313T	(*133*)
1TG2	*Human*	c 117-424	2.2 Å	Fe(III), BH_2, A313T	(*133*)
2V27	*Cp*	4-267	1.5 Å	Fe(III)	(*134*)
2V28	*Cp*	7-267	2.0 Å		(*134*)

As mentioned in section 2.1, the AAAHs are believed to share a similar reaction mechanism. The active site iron in the resting state is 6 coordinated (6C), but upon substrate and cofactor binding the iron coordination changes to 5 coordinated (5C) and large overall structural changes are seen. These two states before and after binding of both substrate and cofactor are referred to as the open and closed conformation, respectively. The structure of the catalytic domain of human TPH1 with BH_2 bound (1MLW)(*47*) represents the open structure with iron being 6C. The substrate binding pocket is accessible and ready for tryptophan to bind. The structure of the catalytic domain of chicken (*gallus gallus*) TPH1 (3E2T)(*12*) on the other hand represents the closed structure. Tryptophan is bound, the binding pocket is closed and only a small opening in the structure gives access to the active site. This structure resembles the structure of the ternary complex of PAH in the closed conformation (1MMK) more than the TPH structure in the open confirmation (1MLW)(*12*). However, in the structure of c*g*TPH1 only tryptophan and no cofactor is bound. An imidazole molecule from the crystallization conditions is coordinated to the active site iron, which is 5C. Although the imidazole is not binding exactly in the cofactor binding pocket, it might mimic the cofactor(*12*), thus inducing the change from the open to the closed confirmation. Another option is that TPH does not follow the general mechanism of PAH and TH and changes from the open to the closed conformation when binding only the substrate. In order to investigate this, a new crystal structure of c*g*TPH1 with tryptophan but without imidazole bound is needed.

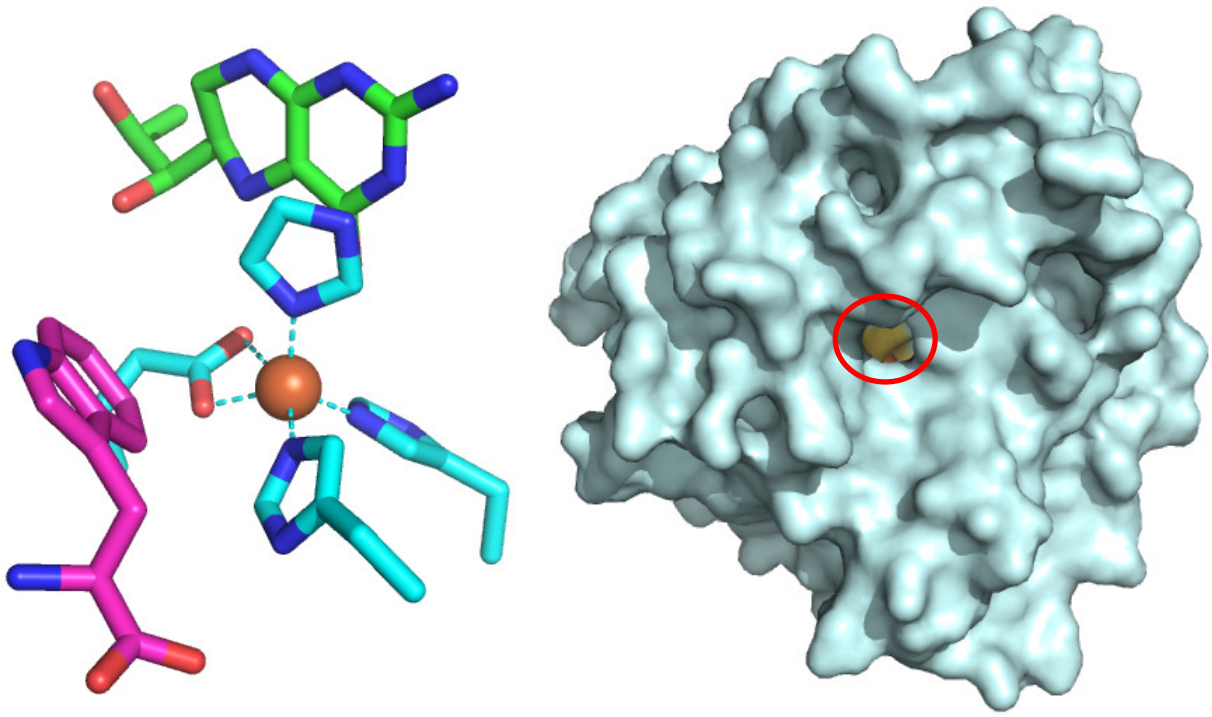


Figure 72: The structure of the catalytic domain of *gallus gallus* TPH1 with tryptophan bound (PDB ID: 3E2T). Left: The iron coordination in 3E2T at the active site. Iron is shown as an orange sphere, BH_4 aligned from the 1MMK structure is shown in green sticks, imidazole is shown in cyan sticks and tryptophan is shown in magenta sticks. Right: Surface representation of the 3E2T structure. The small entrance to the active site is marked with a red ring. Amino acids coordinating iron are colored yellow. The two images are not seen from the same orientation. The figure was made using PyMOL(*50*).

The recently published structures of the catalytic domain of human TPH1 with designed inhibitors bound are all very similar. The amino acid moiety of the inhibitors bind in the active site pocket of tryptophan,

but the inhibitors do not restrict the access to the cofactor binding pocket(*48*). Iron is 6C with 3 waters, a monodentate glutamate and 2 histidines, as in a typical resting state open confirmation for 2 of the inhibitors (PDB ID: 3HF6 and 3HFB). For the last inhibitor (PDB ID: 3HF8), iron is 5C to only one water molecule and featuring a bidentate binding from glutamate. The overall structure is open with access to the active site, even though the tryptophan binding pocket is closed, see Figure 73. Alignment of α C-atoms of 3HF6 with 1MLW (open conformation) and 3E2T (closed conformation) gives a root mean square (RMS) of 0.553 and 0.966, respectively, which supports that the inhibitor bound structure is most similar to the open conformation structure. This is in contrast to the authors opinion in reference (*48*). Most likely the inhibitor structures represent some kind of intermediate between the open and closed conformations.

To obtain a greater knowledge on the structural characteristics of TPH as well as the structure-function relationship, more crystal structures of TPH are needed. Especially structures of isoform 2 or structures containing the tetramerization or regulatory domains would fill in a large gap in the current knowledge. Structures of the catalytic domain with different combinations of tryptophan and/or cofactor bound could also shed light on structural changes compared to PAH. Thus, in this project attempts were made to produce crystal structures of the purified TPH variants. Furthermore, the crystallization of c*g*TPH1 was optimized.

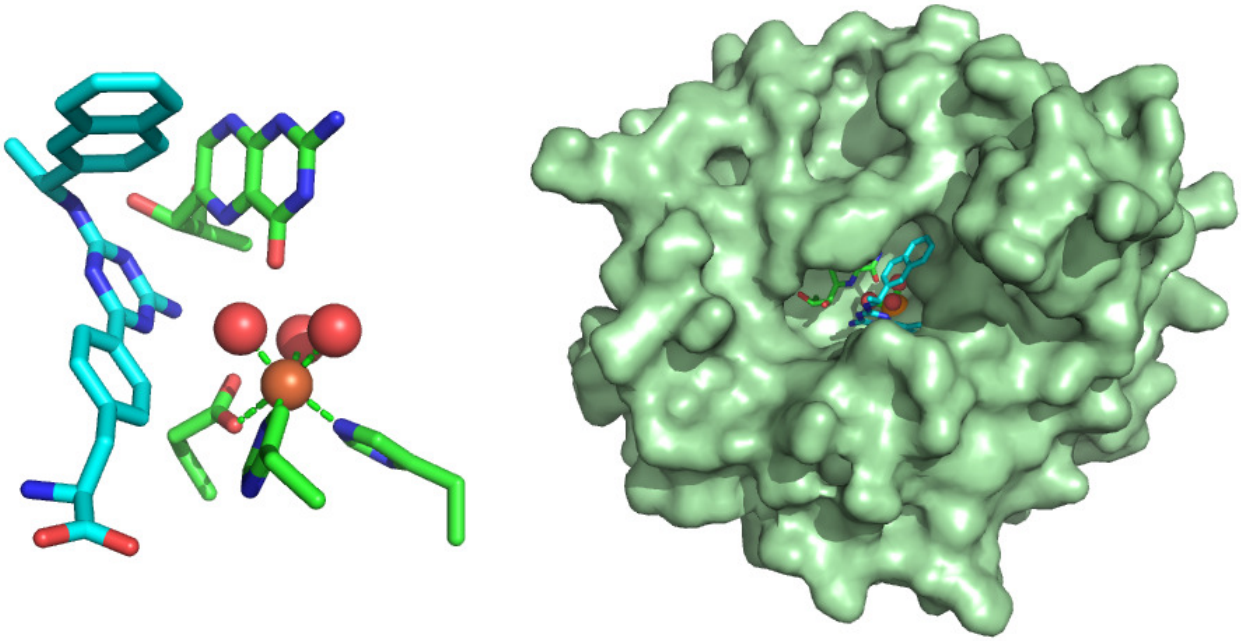


Figure 73: The structure of the catalytic domain of human TPH1 with inhibitor bound (PDB ID: 3HF6). Iron is shown as an orange sphere, water as red spheres, the inhibitor in cyan sticks and BH_2 aligned from the 1MLW structure is shown in green sticks Left: The iron coordination in 3HF6 at the active site. Right: Surface representation of the 3HF6 structure. The active site is open, and both the inhibitor, the cofactor binding pocket (marked by BH_2 in green sticks) and iron coordinating three waters are visible. The two images are not seen from the same orientation. The figure was made using PyMOL(*50*).

5.2 Crystallization of c*g*TPH1

c*g*TPH1 has been purified and crystallized earlier by members of the Metalloprotein Chemistry and Engineering Group(*12*). However, the crystallization procedure was very time-demanding, as it took 3-6 months for the crystals to appear. Furthermore, in the c*g*TPH1 structure imidazole from the crystallization conditions is coordinating to the active site iron. This gives some uncertainties in the analysis of the

structure regarding structural changes upon substrate binding (see section 5.1). In this project the procedure for crystallization of c*g*TPH1 was optimized, so crystals appeared after 3-7 days, and the diffraction quality of the resulting crystals were tested with the goal of solving the structures.

5.2.1 Experimental procedures

c*g*TPH1 was purified as described in reference (*96*), split in two portions, concentrated to 3.3 and 6.3 mg/mL, and frozen at -80 °C until further use. Optimization screens were developed based on the initial crystallization conditions, which were 22.5 (w/v)% PEG 10000, 0.2 M imidazole malate, pH 8.5. Different protein concentrations, addition of substrates, type and concentration of PEG, different additives, as well as temperature variation was tested. For a detailed description of the optimization screens, see appendix 9.2-9.7. Some screens were set up on ice, and stored in a cooled incubator, model 3101 from RUMED set to 5 °C. In general, the sitting-drop trays used were CombiClover Plates from Jena Bioscience, sealed with Crystal Clear sealing tape from Hampton Research. 250 µL screen solution was added into each reservoir. 2 µL protein solution was added in each drop chamber, and was subsequently mixed with 2 µL screen solution from the reservoir by pipetting up and down three times. The drops were inspected using a M205C microscope equipped with a DFC290 camera and the LAS software, v 3.1.0, all from Leica Microsystems.
The crystals were mounted primarily in litho loops from Molecular Dimensions using cryo solutions that were as close to the drop conditions as possible (for exact description of the cryo solutions used for the best diffracting crystals, see section 5.2.3).
The diffraction data were collected by either of two means: I) By testing the crystals ourselves at MaxLab beamline I911-3(*135*). II) By submitting the crystals to the Mail-In service at Brookhaven National Laboratory(*136*). For processing of data images, indexing, cell refinement and integration the software iMosflm 1.0.5(*137*) or XDS(*138*) were used. For molecular replacement and following refinement, CCP4i application suite programs(*139*) Scala, MolRep or Phaser and RefMac were used together with WinCoot(*140*). Molecular replacement was performed with the 3E2T structure where everything except for the amino acid chain had been removed. Alignment between different structures were performed using the align command in PyMOL(*50*).

5.2.2 Crystallization results

The initial crystallization conditions for c*g*TPH1 were: 22.5 (w/v)% PEG 10000, 0.2 M imidazole malate, pH 8.5 and 2.8 mg/ml (78 µM) c*g*TPH1 stored at 4 °C(*12*). No substrates were added. The structure contained tryptophan and imidazole bound in the active site. As the crystallization conditions for c*g*TPH1 were already known, only the fine-tuning of parameters was investigated.

5.2.2.1 1st optimization screen

First, an optimization screen based on the initial crystallization conditions of c*g*TPH1 was made and set up at room temperature and 5 °C. The screen varied protein concentration and PEG type and concentration. The conditions used in the screen can be seen in appendix 9.2. Tryptophan and BH_2 were added both individually and combined. BH_2 is added instead of the natural cofactor BH_4, to avoid the reaction occuring, and because BH_4 is not stable at room temperature at neutral pH. Tryptophan and BH_2 was added in at least 2 fold molar excess from 7 mM stock solutions. Crystals appeared in the screen set up at room temperature after only 7 days, but only in drops containing imidazole and added tryptophan or tryptophan and BH_2. The crystals all grew from precipitate. Most of the crystals seemed flat rather than 3 dimensional, but some conditions gave very nice crystals. The screen at 5 °C only produced crystals in the drops with

conditions similar to the original, but with added tryptophan. These crystals appeared after 14 days and looked very three dimensional. Thus, it can be concluded that adding tryptophan to the crystallization conditions decreases the crystallization time significantly, and that incubation of the screen at room temperature speeds up the crystallization process even more. Some of the crystals from the first screen can be seen in Figure 74. Several crystals were tested, but none gave diffraction data good enough for structure solving as the resolution was too low (4-8 Å).

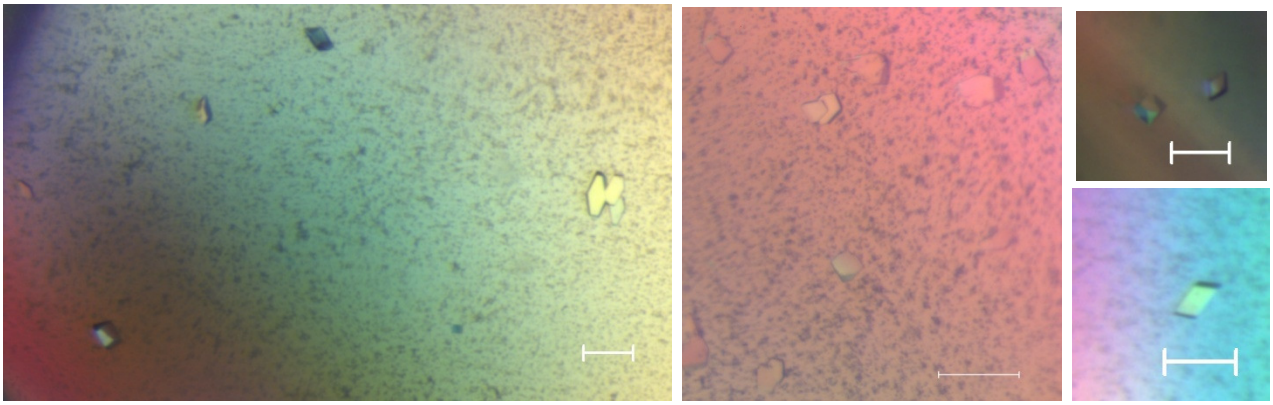


Figure 74: Pictures of *cg*TPH1 crystals from the 1st optimization screen. The left picture is from the screen at 5 °C, while the other 3 pictures are from the screen at room temperature. The white scale bar in the bottom right corner of each picture is 0.1 mm.

5.2.2.2 2nd optimization screen

To check if any additives would speed up the crystallization process, an additive screen from Molecular Dimenions was tested with the original crystallization conditions (22.5 (w/v)% PEG 10000, 0.2 M imidazole malate pH 8.5) and a protein concentration of 2.8 mg/mL at room temperature. The additives were added according to the screen application note. Tryptophan and BH_2 were added both individually and combined. A description of the additive screen can be found in appendix 9.3. No crystals appeared in this screen.

5.2.2.3 3rd optimization screen

In the 1st optimization screen, only conditions with imidazole gave crystals. In order to obtain crystals from conditions without imidazole, a new optimization screen was set up. The screen can be seen in appendix 9.4 and was set up at room temperature.

Crystals appeared after only 3 days in several conditions, which all had been added tryptophan or tryptophan and BH_2. 7 conditions without imidazole gave crystals. The drops mainly contained many small crystals, but some larger crystals also appeared. Several crystals were tested with X-ray diffraction, but only 3 gave diffraction data good enough for structure solving (see section 5.2.3.1 and 5.2.3.2). Most of the crystals gave only low resolution data (4-8 Å). The 3 crystals that diffracted to high enough resolution for structure solving all came from this screen. Furthermore, it can be concluded that c*g*TPH1 crystals can be grown without imidazole in the crystallization conditions.

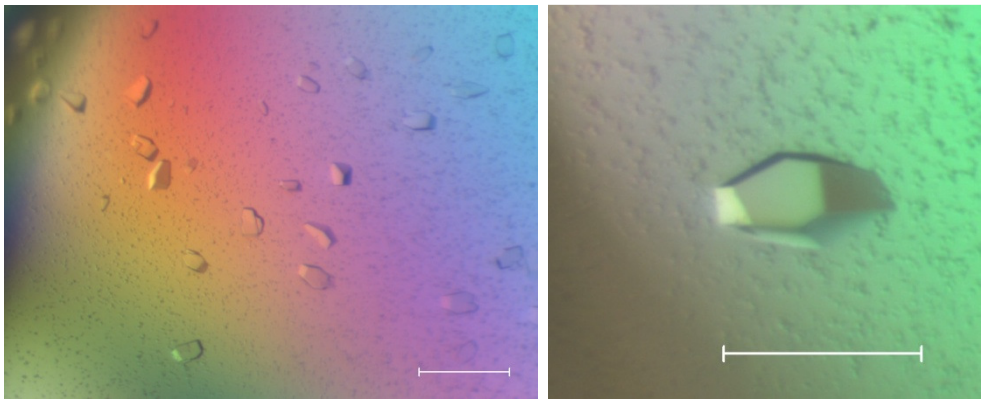


Figure 75: Pictures of *cg*TPH1 crystals from the 3rd optimization screen. The white scale bar in the bottom right corner of each picture is 0.1 mm.

5.2.2.4 4th optimization screen

In an attempt to get larger and better quality crystals without imidazole, a 4th optimization screen was set up, based on the previous results. The screens can be seen in appendix 9.5, 9.6, and 9.7. Larger drop volumes and different kind of crystallization trays/plates including hanging drops were tested in order to get larger crystals. Tryptophan or tryptophan and BH_2 were added to all conditions, and the screen was set up at room temperature.

Approx. half (38) of the conditions tested gave crystals, many of which looked suitable for structure determination in terms of size and three dimensional shape. Especially, the screen with hanging drops produced many crystals. The crystals appeared after 3-4 days. Many of the crystals from this screen were sent for analysis at Brookhaven National Laboratory, but unfortunately none gave data of high enough quality for structure determination (resolution around 4 Å). It can be speculated that the shorter crystallization times gives crystals of poorer quality compared to the original crystals, which were grown over 3-6 months. The literature only gives sparse information on the correlation between crystal growth rate and crystal quality(*141*, *142*). However, it seems reasonable that faster crystal growth rates would increase the amount of defects incorporated into the crystal.

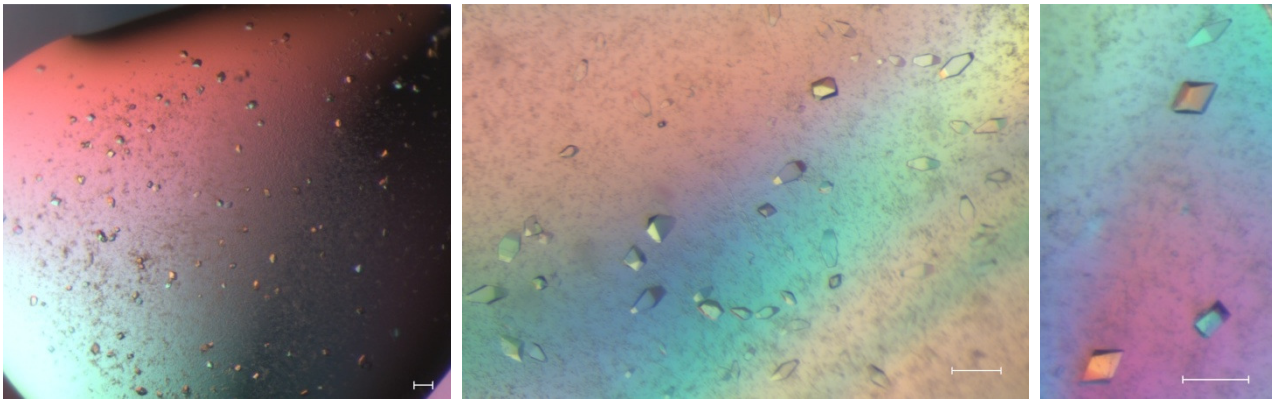


Figure 76: Pictures of *cg*TPH1 crystals from the 4th optimization screen. The white scale bar in the bottom right corner of each picture is 0.1 mm.

5.2.3 Structures of c*g*TPH1

Of the many crystals produced, several were tested by X-ray diffraction. Most crystals gave low resolution data (4-8 Å), which could not be used to solve the structure. 30 crystals were shipped to Brookhaven National Laboratory for analysis, but of these 30 crystals only 2 diffracted well enough for structure solving. These two crystals were the only two that contained imidazole in the crystallization conditions. Based on this, it seems that imidazole affects the crystallization and helps creating crystals that gives high resolution diffraction data. One crystal without imidazole was measured at MaxLab in Lund, Sweden and gave diffraction data with high enough resolution for structure solving. The two structures are presented in the following two sections.

It should be noted here that there was no correlation between the crystal appearance and diffraction data quality. Many large and edgy crystals gave low resolution data, while small and seemingly flat crystals resulted in the best diffraction data.

5.2.3.1 Structure of cgTPH1 with no imidazole and tryptophan bound

The first crystal of c*g*TPH1 which gave reasonable diffraction data was obtained from the 3rd optimization screen. It will be referred to as crystal-1 and grew in the following conditions: 22.5 (w/v)% PEG 10000, 0.1 M Tris acetate pH 8.5. The protein concentration was 2.5 mg/ml and tryptophan was added in 4.5 times excess (310 µM) compared to c*g*TPH1 (70 µM). The crystals appeared after 3 days at room temperature.

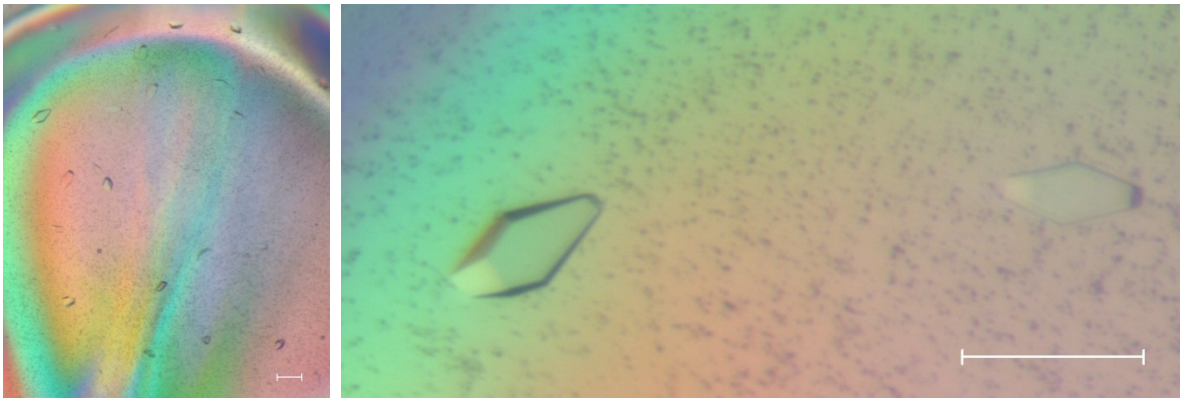


Figure 77: Crystal-1 (c*g*TPH1) from the 3rd optimization screen. The white scale bar in the bottom right corner of each picture is 0.1 mm.

The crystal was mounted and frozen in cryo-solution containing 22.5 (w/v)% PEG 10000, 20 (v/v)% PEG 400, 0.1 M Tris acetate pH 8.5, 155 µM tryptophan. The diffraction data was collected at MaxLab, Lund, Sweden at beamline I911-3. The results of the diffraction data collection can be seen in Table 13.

Because of the low quality data, only the overall structure was solved and the resolution was limited to 3.1 Å. No attempts were made to refine the structure. From the initial map, electron density was seen in the position where iron usually is placed. Thus iron was added to the structure. However, on addition of iron the R_{Free} factor increased, even though the R factor decreased, see Table 14. This effect is usually seen when the structure is over-refined(*143*).

Table 13: Parameters from the data collection of crystal-1 (c*g*TPH1) at MaxLab, Lund, Sweden.

Space group	C121
X-ray source	I911-3, MAXII, MaxLab Lund
Detector	Marmosaic 225
Wavelength (Å)	1.00
Data collection temperature (K)	100
Resolution limit (Å)	47.7-2.5
No. reflections	51678
Unique reflections	19336
Redundancy	2.7
Completeness (%)	93.8
R_{merge}	0.297
<I/σ(I)>	2.5
Unit cell parameters **a (Å), b (Å), c (Å)** **α (°), β (°), γ (°)**	 171.5, 49.7, 67.6 90.0, 91.9, 90.0
Solvent content	38.4
Matthews coefficient ($Å^3$/Da)	1.99
Nmol/assym.	2
Average mosaicity	1.04

The new electron density map then revealed excess electron density in the tryptophan binding pocket and thus tryptophan was added in this position. The electron density map of the active site from the final refinement can be seen in Figure 78. It should be noted that R_{Free} increased for both addition of iron and tryptophan. An excess electron density blob in the coordination sphere of iron was seen in some of the first refinement rounds, however, when water was added it moved far away from iron during the following refinement (approx. 6 Å). Thus, no conclusions can be made on the complete coordination of iron.

Table 14: Overview of the R and R_{Free} values upon refinement of the structure of crystal-1 (c*g*TPH1 with tryptophan but no imidazole bound).

Operation	R	R_{Free}
Rigid body refinement	0.3686	0.3630
Restrained refinement	0.2587	0.3391
Addition of iron	0.2456	0.3463
Addition of tryptophan	0.2385	0.3558

The overall fold of this newly solved structure is more similar to the closed confirmation of 3E2T than to the open confirmation of 1MLW. However, 3E2T was also used as search model during molecular replacement. Furthermore, the data are of such low quality and the R and R_{Free} values are correspondingly high and separated by too many percent, to make any definite conclusions from this structure.
In conclusion, the crystal did contain c*g*TPH1, but the diffraction data were not good enough to make any definitive conclusions on the overall or active site structure.

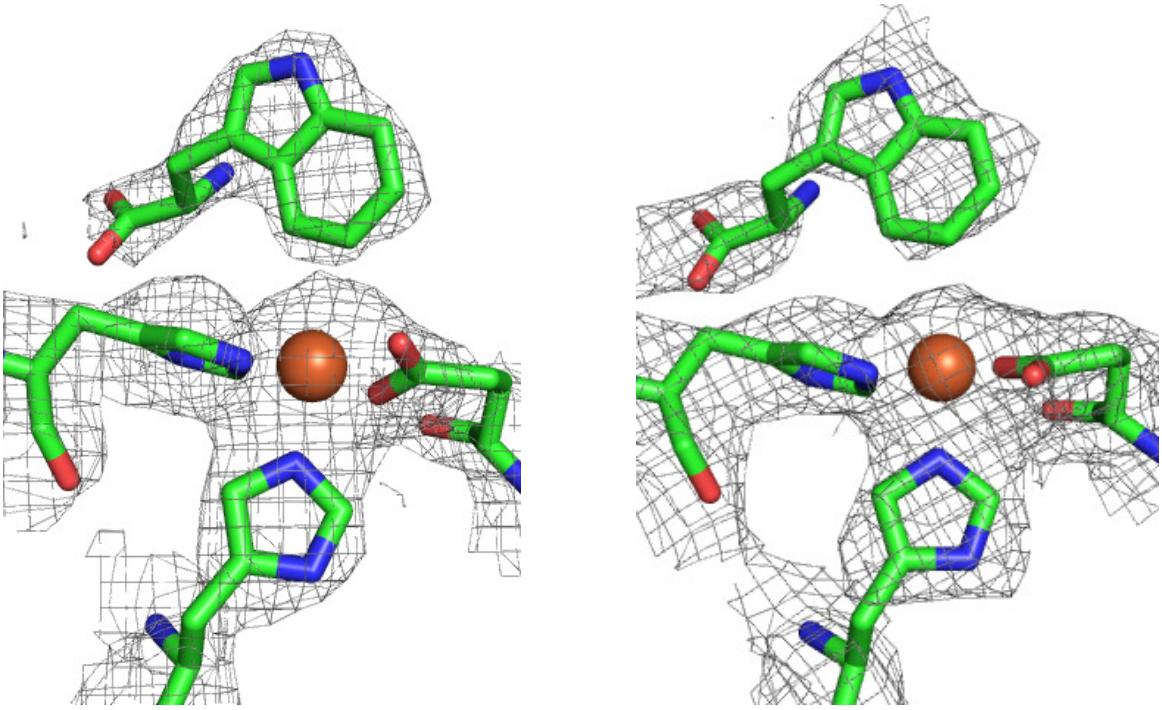


Figure 78: Electron density map of the active sites in the 2 molecules in the asymmetric unit of the structure of crystal-1 (c*g*TPH1 with tryptophan but with no imidazole bound). Iron is shown as an orange sphere. σ=1 for the electron density. The figure was made using PyMOL(*50*).

5.2.3.2 Structure of cgTPH1 with imidazole and tryptophan

Brookhaven National Laboratory Mail-in service measured 30 c*g*TPH1 crystals from the 3rd and 4th optimization screen, and of those only 2 diffracted well. These two crystals, referred to as crystal-2 and crystal-3, came from the same drop from the 3rd optimization screen containing 22.5 (w/v)% PEG 10000, 0.2 M imidazole malate pH 8.5. The protein concentration was 2.8 mg/ml and tryptophan and BH_2 was added in 2 times excess (156 µM) compared to c*g*TPH1 (78 µM). The crystals were analyzed because it was hoped that both tryptophan and BH_2 would have bound in the active site, and the structure thus would represent a ternary complex. The crystals had a different shape (triangular) compared to crystals with only tryptophan added (rhombic), and can be seen in Figure 79. The crystals appeared after 3 days at room temperature.

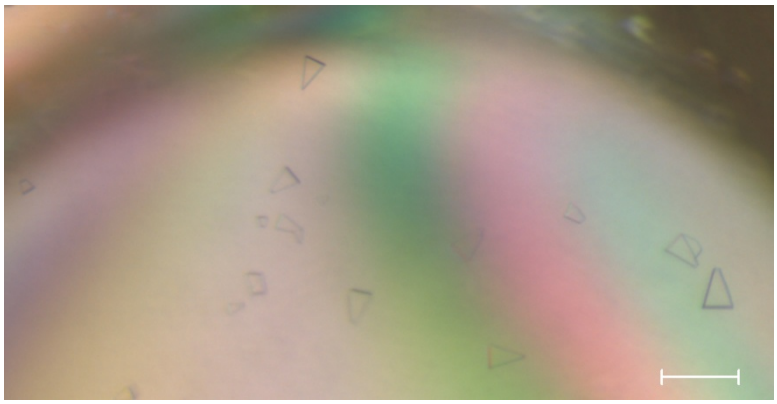


Figure 79: Crystal-2 and crystal-3 (cgTPH1) from the 3rd optimization screen. The white scale bar in the bottom right corner of the picture is 0.1 mm.

The crystals were mounted and frozen in a cryo-solution containing 22.5 (w/v)% PEG 10000, 20 (v/v)% PEG 400, 0.1 M imidazole malate pH 8.5, 156 µM tryptophan, 156 µM BH_2. All measurements and data analysis including molecular replacement with 3E2T as search model and the following refinement was performed by the Mail-in service crew at Brookhaven National Laboratory and will not be described further. The parameters for the two structures can be seen in Table 15. The two structures were very similar, and thus only the one structure of crystal-2 will be presented here.

Table 15: Parameters from the data collection of crystal-2 and crystal-3 of c*g*TPH1 at Brookhaven National Laboratory mail-in service, New York, USA.

Parameter	Crystal-2	Crystal-3
Space group	C121	C121
Resolution limit (Å)	2.4	2.3
Unit cell parameters **a (Å), b (Å), c (Å)** **α (°), β (°), γ (°)**	 170.5, 48.7, 135.1 90.0, 91.7, 90.0	 170.7, 48.7, 134.9 90.0, 91.6, 90.0
Nmol/assym.	4	4
No. of reflections	34324	40815
Completeness	81.8	85.5
R_{Free}	0.27	0.25
R	0.19	0.19

The structure of crystal-2 was refined to R_{Free} = 27 %. Iron and tryptophan are bound in the active site in the structure. An imidazole ring has also been added near the iron coordination sphere. From the electron density map in Figure 80 it is seen that a ring-like molecule binds in the vicinity of iron, but it is not possible to determine if a water molecule is also coordinated to iron.

Imidazole is located near the cofactor binding pocket, but the imidazole ring does not overlay the rings of the cofactor in the 1MLW structure or 1MMK structure. Furthermore, the electron density does not allow for a precise positioning of imidazole, which differs somewhat in the 4 molecules in the asymmetric unit. Thus, it is hard to interpret whether imidazole mimics the cofactor. As seen by the electron density maps, tryptophan is clearly bound in the structure, and both tryptophan and iron are present with 100 % occupancy. In Table 16, the distances between the different molecules in the structure of crystal-2 are seen. It can be discussed if imidazole coordinates to iron, considering the long distances between the two.

The overall fold of the structure is very similar to that of the previous c*g*TPH1 structure 3E2T, also with imidazole and tryptophan bound. The RMS value only aligning α carbon atoms is 0.3 for 3E2T and 1.3 for 1MLW. The alignment with the ternary complex of PAH (PDB ID: 1MMK) gave an RMS of 0.6, suggesting that the structure is in the closed conformation. However, this structure does unfortunately not contribute with any new structural information on TPH.

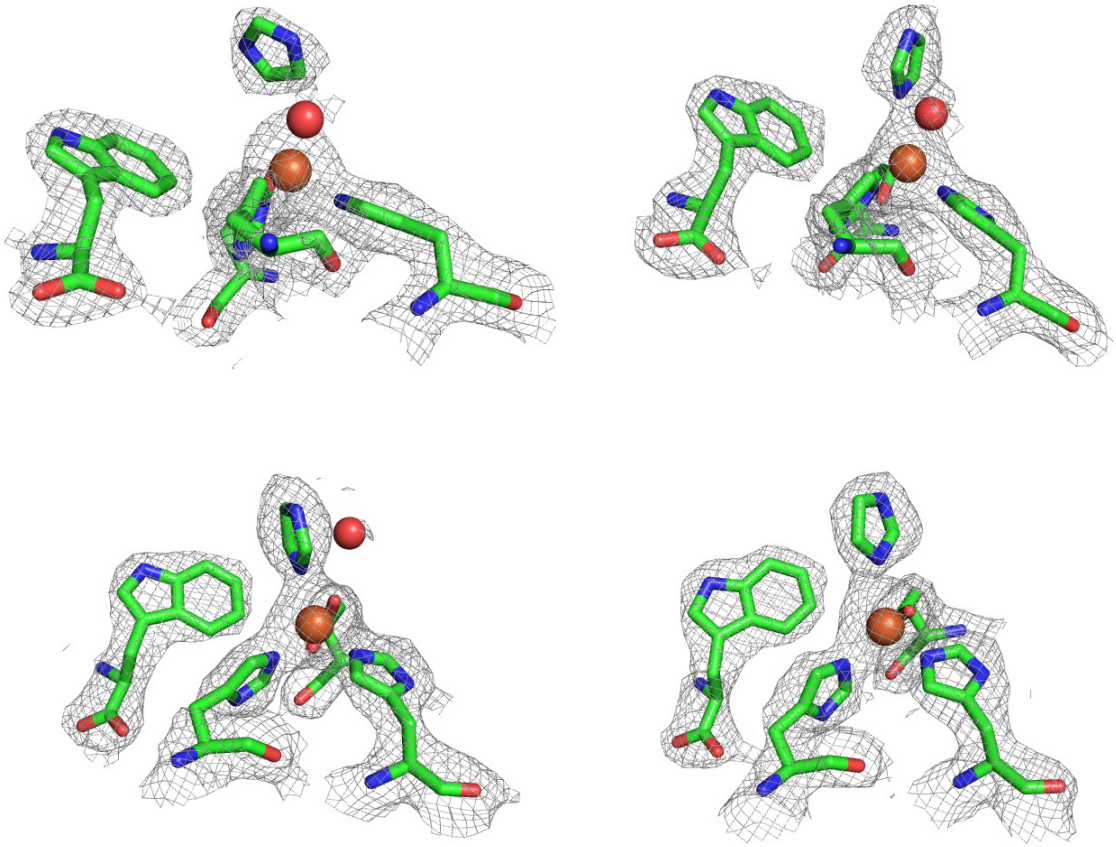


Figure 80: Electron density map of the active sites in the 4 molecules in the asymmetric unit of crystal-2 (cgTPH1 with tryptophan and imidazole bound). Iron is shown as an orange sphere and water as red spheres. σ=1 for the electron density. The figure was made using PyMOL(*50*).

Table 16: Distances between different molecules in the structure of crystal-2 (c*g*TPH1 with tryptophan and imidazole bound). The values are given for each of the 4 molecules in the asymmetric unit.

Distance	Molecule 1 [Å]	Molecule 2 [Å]	Molecule 3 [Å]	Molecule 4 [Å]
Fe -> imidazole	3.6	2.8	2.3	2.7
Fe -> water	2.2	2.5	3.7	-
Fe -> tryptophan	4.1	4.0	3.9	4.0
Fe -> His278	2.0	1.7	2.0	2.0
Fe -> His273	2.1	2.1	2.0	2.1
Fe -> Glu318	1.7 (2.8)	1.9 (2.4)	1.9 (2.5)	1.6 (2.8)

5.2.4 Conclusions

The crystallization procedure for c*g*TPH1 has been optimized to shorten the crystal growth period significantly from 3-6 months to a few days by addition of tryptophan and incubation at room temperature. Nice crystals of c*g*TPH1 can also be grown without imidazole in the crystallization conditions. Even though the c*g*TPH1 crystals all looked very edged and three dimensional, the diffraction data were of poor quality. The best diffraction data came from crystals grown with imidazole, but the solved structures did not contribute with any new knowledge on the structure of c*g*TPH1.

5.3 Crystallization of c*h*TPH2

One of the main goals in this project was to obtain a structure of TPH2. Thus, large efforts were put into the attempts to crystallize c*h*TPH2 and ct*h*TPH2. Crystallization of c*h*TPH2 was attempted using several different buffer systems for purification, different screens, protein concentrations, addition of substrates, and temperature variation. A total number of approx. 4000 drops with c*h*TPH2 were set up manually. Crystals did appear, but they turned out to be organic molecule crystals, possibly from the buffer. Protein samples of c*h*TPH2 were shipped to HT (high throughput) crystallization at EMBL in Hamburg through the P-CUBE project(*144*). Here, 5 different screens of 96 conditions were tested with 3 different protein concentrations and 4 combinations of substrates added at 19 °C to a total of 5760 drops. No crystals were found.

In my master thesis, experiments with different scattering techniques showed that purified c*h*TPH2 is not completely monodisperse, and that part of the sample might be denatured or exist as dimers(*11*). On SDS-PAGE the sample looks pure (see section 4.3.3.1). Despite the multiple attempts to crystallize c*h*TPH2 no crystals were made. Therefore, a more strategic way for testing of the samples and crystallization was chosen (see section 5.5).

5.4 Crystallization of ct*h*TPH2

Earlier projects investigated the monodispersity of a ct*h*TPH2 solution, and found that it was monodisperse and that ct*h*TPH2 exists as tetramers(*11*). ct*h*TPH2 purified in this project also looked pure judged from the SDS-PAGE (see section 4.3.3.2). As with c*h*TPH2 several different buffer systems for purification, different screens, protein concentrations, addition of substrates, and temperature variation were tested in a total of approx. 5000 drops. The only crystals that appeared turned out to be organic molecule crystals, most likely of BH_2 or buffer.

5.5 Crystallization pipeline

Along the project course it was clear that several new variants of TPH could be purified and was thus ready for crystallization trials. Because of the many different variants of TPH and the unsuccessful crystallization trials of c*h*TPH2 and ct*h*TPH2, a more strategic and systematic way of analyzing the purified samples and testing for crystallization was set up, called the crystallization pipeline. Even though the strategy for reaching the desired results was changed, the overall goal remains the same: To crystallize and solve the structures for TPH variants where the structure is unknown or known structures in new combinations with substrate and cofactor.

The TPH variants that can be purified in reasonable amounts and are thus available for crystallization are shown in Table 17.

These 5 variants constitute the main candidates in the crystallization pipeline in this project. c*g*TPH1 has not been included as it has already been crystallized. c*h*TPH1 has been included even though it has been crystallized in the group before(*10*) because an impurity has been removed in the new purification method (see section 4.3.2.1), and this might improve the crystallization process and crystal quality.

Table 17: Overview of TPH variants that can be purified in reasonable amounts suitable for crystallization. Concentration is the highest concentration of the protein tested without precipitation. The same data is given in Table 11.

Variant	Purification method	Buffer	Yield	Concentration	Section
c*h*TPH2	Anion-exchange and gel filtration	20 mM HEPES/NaOH 100 mM $(NH_4)_2SO_4$ pH 7.2	70 mg/L culture	2.5 mM	4.3.3.1
ct*h*TPH2	Anion-exchange and gel-filtration	20 mM HEPES/NaOH 100 mM $(NH_4)_2SO_4$ pH 7.2	54 mg/L culture	0.29 mM	4.3.3.2
c*h*TPH1	Affinity using GST-tag and gel-filtration	1xPBS 200 mM $(NH_4)_2SO_4$ pH 8.5	11 mg/L culture	0.18 mM	4.3.2.1
rc*h*TPH1	Affinity using MBP-tag and gel-filtration	20mM Tris/H_2SO_4 300mM $(NH_4)_2SO_4$ pH 8.0	20 mg/L culture	0.20 mM	4.3.2.4
***h*TPH1**	Affinity using MBP-tag and gel-filtration	20 mM Tris/H_2SO_4 300 mM $(NH_4)_2SO_4$ pH 8.0	23 mg/L culture	0.16 mM	4.3.2.2

5.5.1 Strategic approach

When a TPH variant can be purified, the purification product should be tested in order to ensure that the protein sample is pure, active, and stable. This is tested through SDS-PAGE, mass spectrometry, activity assay, and DSF measurements. The monodispersity of the sample after freezing can be tested using analytical gel-filtration. A pure and uniform sample is one of the most important parameters in crystallization(*95*). If these tests show a pure, active, stable, and monodisperse protein sample, it should be suitable for crystallization.

The monodispersity of the different TPH variants was tested in chapter 4, and an overview of the results is seen in Table 18. Most samples are slightly polydisperse after freezing and thawing, and this could explain the difficulties with crystallization. The aggregates or unfolded proteins represented by the shoulder on the main peak in the chromatograms were not separated by centrifugation or filtration as they are too small and soluble. Whether it is the freezing/thawing or the concentration process that causes the low monodispersity is unknown. However, most samples gave chromatograms with only very small shoulders on the main peak, which indicates that by far the majority of the sample has the same size. For crystallization purposes, even small amounts of unfolded or aggregated protein can inhibit or complicate the process(*95*, *122*). Thus, if crystallization fails, the monodispersity of the sample would be one of the first parameter that should be investigated and possibly improved. However, as aggregates in some cases might abet crystallization, centrifugation or filtration of a sample is generally not advised prior to crystallization(*125*). It can thus be discussed if a slightly polydisperse sample is good or bad for crystallization. Screening of a large number of proteins has shown that a symmetric gel-filtration chromatography profile in general seems to be required for successful crystallization(*145*). However, most gel-filtration chromatograms in the reference which resulted in crystal structures also displayed small shoulders near the main peak as most TPH variants. Known protein stabilizers such as glycerol might help increasing the monodispersity, however they might also interfere with the crystallization process or even

bind to the active site (*vide infra*, see section 6.3). For ct*h*TPH2, the presence of 10 (w/v)% glycerol in the buffers did not improve the monodispersity (see section 4.3.3.2.1). To eliminate the possible negative effect of freezing and thawing, the TPH variants were purified and the screens for crystallization were set up directly after purification.

Table 18: Overview of the characterization results from the purification products of each variant. Details about the individual results can be seen in chapter 4.

Variant	SDS-PAGE	Mass spectrometry	Activity assay	T_M from DSF at pH 7.5	Monodispersity after freezing
c*h*TPH2	Pure	Pure monomer with iron bound	Active	53	Small shoulder that cannot be removed
ct*h*TPH2	Pure	Pure tetramer	Active	56	Small shoulder that cannot be removed
c*h*TPH1	Pure	Pure monomer with iron bound	Active	46	Small shoulder that can be removed
rc*h*TPH1	Pure	Mixture of several oligomer species	Active	-	Monodisperse
***h*TPH1**	Pure	Pure tetramer	Active	54	Large shoulder that can be decreased but not completely removed

5.5.2 Screens

When the TPH variant has been purified and the purification product has been tested and found suitable for crystallization, the actual crystallization set-up process can begin. It was decided to test all variants at the same conditions to ensure a better comparison between variants, and to limit the number of variables. The sparse matrix screen Structure Screen 1 from Molecular dimensions was chosen as the main screen. Conditions containing calcium (condition 1, 11, 20, and 23) was removed, as the protein buffers contain sulfate which would crystallize with calcium as $CaSO_4$. Instead, the crystallization conditions for c*g*TPH1 and condition 26 from Structure screen 2 similar to the crystallization conditions for c*h*TPH1 were included, giving a total of 48 conditions. These 48 conditions were tested with 2 protein concentrations (2 mg/mL and 5 mg/mL) and either no addition of substrates or addition of both tryptophan and BH_2 in 3 fold molar excess. One screen was set up at room temperature and the other at 5 °C. This gives a total of 384 drops for each TPH variant. Each drop contains 2 µl protein solution and 2 µl reservoir solution. Thus, 2.7 mg protein was used for the entire experiment for each variant.

It was considered to add iron to the crystallization conditions. However, iron is very hard to keep soluble at pH above 5 when oxygen is present, as it precipitates as Fe(III) oxides. Thus, Fe(II) was added during purification and not in the crystallization conditions.

5.5.3 Experimental procedures

The experimental methods used in the crystallization pipeline will be presented in this section, divided into the protein sample preparation and the screen set-up procedure.

5.5.3.1 Protein sample preparation

c*h*TPH2, ct*h*TPH2, c*h*TPH1 (from GST-c*h*TPH1), rc*h*TPH1 (from MBP-rc*h*TPH1) and *h*TPH1 (from MBP-*h*TPH1) were purified as described in sections 4.2.4.1, 4.2.4.2, 4.2.3.1, 4.2.3.3, and 4.2.3.2, respectively. For each

variant, the following preparations were made. The final purification product from the last gel-filtration was concentrated using ultrafiltration to a concentration of 7-10 mg/mL or until precipitation was observed. The pressure was taken of the ultrafiltration cell and the solution was left for 15 min to allow the equilibrium between precipitated and soluble protein to adjust. The protein solution was centrifuged at 16000 x g for 15 min at 5 °C. The supernatant was collected and the concentration was determined by UV-Vis. At least three consecutive concentration measurements were performed. Frozen stocks of 7 mM BH_2 and 20 mM L-tryptophan were thawed. Each protein solution was divided into 4 different solutions for crystallization set-up by diluting in the gel-filtration buffer from the purification (see Table 17), and tryptophan and BH_2 were added in 3 times molar excess. An overview of the 4 solutions can be seen in Table 19. The protein solutions were kept on ice during the entire process.

Table 19: Overview of the 4 protein solutions for crystallization set-up that were made for each TPH variant.

Solution	A	B	C	D
Protein concentration	5 mg/mL	5 mg/mL	2 mg/mL	2 mg/mL
Added substrates	-	BH_2+trp	-	BH_2+trp

5.5.3.2 Screen set-up

The screen used was the Structure Screen 1 from molecular dimensions, see appendix 9.8. Conditions 1, 11, 20, and 23 were not set-up, and instead a premade solution of 22.5 (w/v)% PEG 10000, 0.2 M imidazole malate pH 8.5 from Molecular Dimensions and condition 26 from Structure Screen 2 containing 0.2 M $(NH_4)_2SO_4$, 0.1 M MES (2-(N-morpholino)-ethanesulfonic acid), pH 6.5, 30 (w/v)% PEG 5000 MME (monomethylether) were included. The screen solutions were approx. 5 °C when setting up the 5 °C screen, but were allowed to equilibrate to room temperature as the room temperature screen was set up. For the 5 °C screen, the crystallization plate was kept on a tray filled with ice during set-up, and afterwards it was transferred to a cooled incubator, model 3101 from RUMED set to 5 °C where it was stored.

The sitting-drop trays used were CombiClover Plates from Jena Bioscience, sealed with Crystal Clear sealing tape from Hampton Research. 250 µL screen solution was added into each reservoir. 2 µL protein solution was added in each drop chamber, and was subsequently mixed with 2 µL screen solution from the reservoir by pipetting up and down three times.

The drops were inspected using a M205C microscope equipped with a DFC290 camera and the LAS software, v 3.1.0, all from Leica Microsystems.

5.5.4 Results and discussion

In general, 80-90 % of the drops contained protein precipitate. This was common for all TPH variants, despite the relatively low protein concentrations used (2 and 5 mg/mL). However, crystals can also grow from precipitate as in the case for e.g. c*g*TPH1.

The pipeline screening of c*h*TPH2, ct*h*TPH2, rc*h*TPH1, and *h*TPH1 did not result in any crystals. However, for rc*h*TPH1 crystalline precipitate as well as assemblies of precipitate which also appeared crystalline was observed in several conditions. Examples of this are shown in Figure 81. These leads should be optimized to check whether crystals can be obtained.

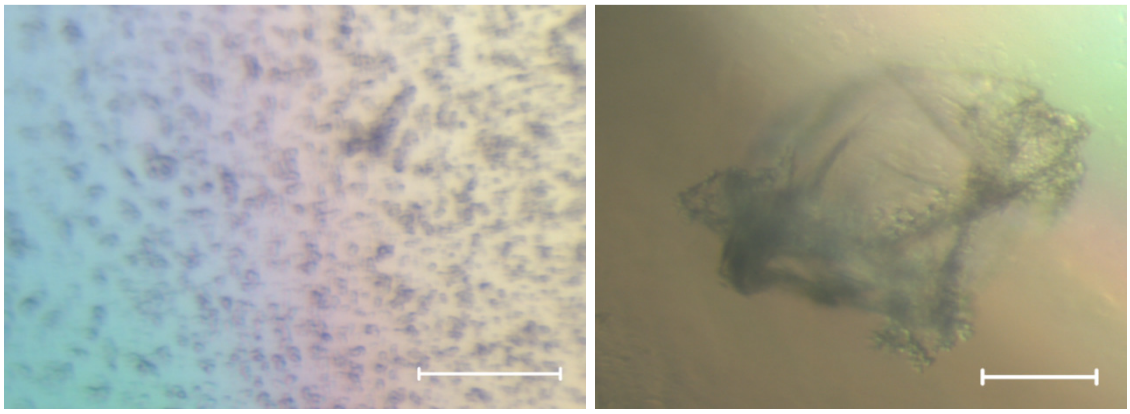


Figure 81: Examples of leads from the pipeline screening of r*ch*TPH1. Left is what is believed to be crystalline precipipate, and right is an assembly of precipitate. The crystals appeared in screen conditions 30 and 26, respectively from left to right (see appendix 9.8 for exact conditions). The white scale bar in the bottom right corner of each picture is 0.1 mm.

For c*h*TPH1, several conditions yielded crystals. A total of 20 different screen conditions successfully produced crystals, while approx. half of them had more than 3 hits (meaning that at least 3 of the 8 drops set up with this condition produced crystals). Most crystals were needle-shaped, as has been observed for c*h*TPH1 before(*10*). Due to the many nucleation points in these drops, the crystals will be hard to separate. Furthermore, several crystals often grew from the same nucleation point in a star like pattern.

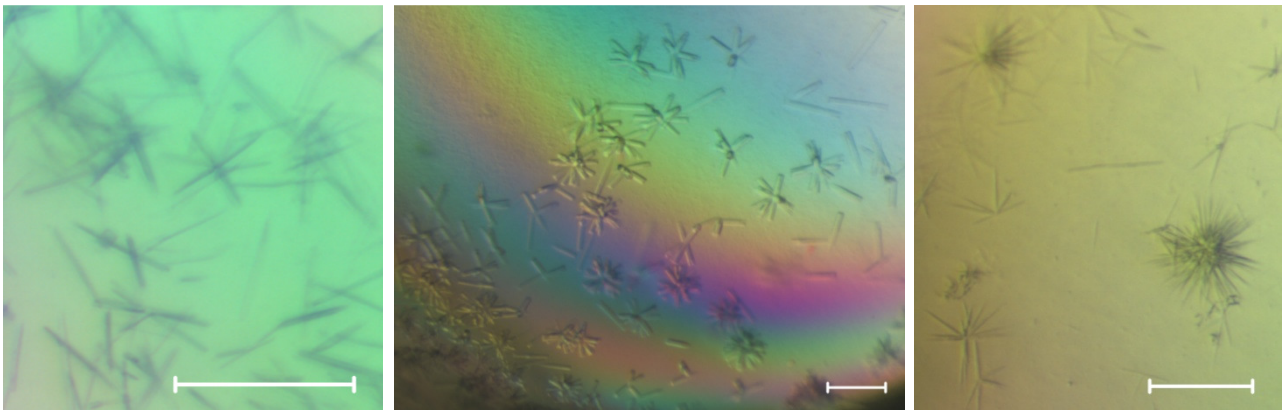


Figure 82: Needle-like crystals from the c*h*TPH1 pipeline screening. The crystals appeared in screen condition 18, 36 and 7, respectively from left to right (see appendix 9.8 for exact conditions). The white scale bar in the bottom right corner of each picture is 0.1 mm.

As the needles are most likely protein crystals, they offer a great starting point for fine-tuning and optimization to yield separate crystals that can be analyzed by diffraction. The number of nucleation points should be decreased, so larger crystals can grow. There are several strategies to overcome the problem of excess nucleation resulting in the formation of numerous low-quality crystals. One option is seeding with already-grown small crystals into a drop in the meta-stable zone, which facilitates crystal growth but not nucleation(*125*).

Several other crystal types than needles were also observed, including spherulites, see Figure 83. However, some of these could be suspected to be salt crystals, due to the intense color when viewed through a polarizer(*122*). As the protein solution contains quite high salt concentrations, it is not unlikely that salt crystals will appear.

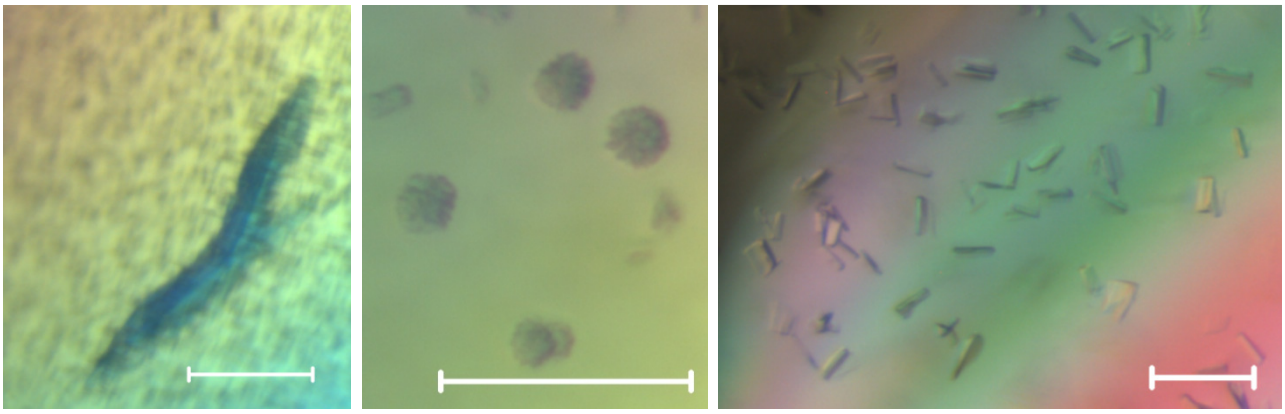


Figure 83: Different types of crystals observed in the *ch*TPH1 pipeline screening. The center picture shows spherulites. The crystals appeared in condition 22, 37 and 35, respectively from left to right (see appendix 9.8 for exact conditions). The white scale bar in the bottom right corner of each picture is 0.1 mm.

Finally, two different conditions (16 and 24) produced very nice looking, edgy and three dimensional rectangular crystals, as seen in Figure 84. Again, the intense color of the crystals when viewed through the polarizer indicates that they might be salt crystals. However, the crystals should be mounted and tested by diffraction to confirm this. Alternatively, the drops can be added Izit crystal dye, which mainly colors protein crystals, and not salt crystals or it can be tested if a mixture of protein-free buffer and screen solution also result in crystals. As the crystals appeared shortly before the deadline of this project, none of the above tests were performed.

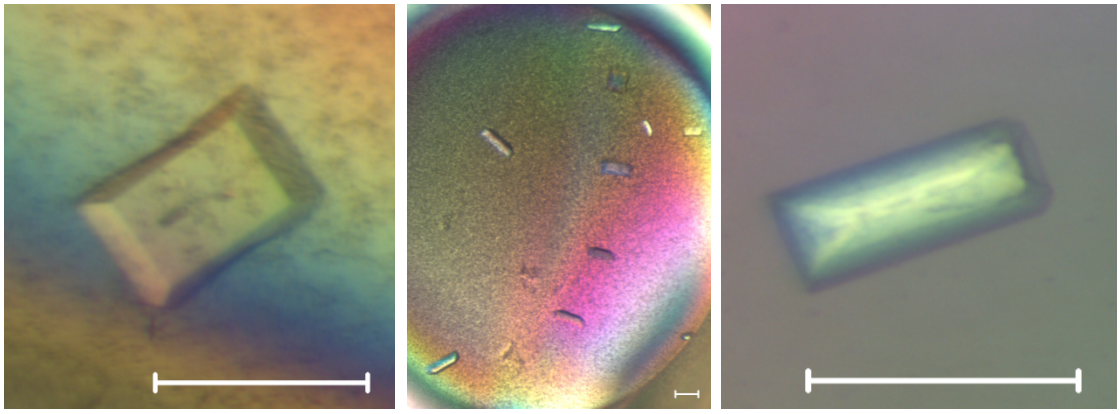


Figure 84: Rectangular crystals observed in the *ch*TH1 pipeline screening. The crystals appeared in condition 16 (two first pictures) and 24, respectively from left to right (see appendix 9.8 for exact conditions). The white scale bar in the bottom right corner of each picture is 0.1 mm.

5.6 Conclusions

The crystallization procedure for *cg*TPH1 was optimized to faster crystal growth by addition of tryptophan and incubation at room temperature. Crystals without imidazole in the crystallization conditions could be obtained. Only one such crystal and two crystals with imidazole gave diffraction data good enough for structure solving. There was no correlation between crystal appearance and diffraction data quality, as the smallest and most flat-looking crystals gave the best diffraction data.

A strategic approach to crystallization of TPH variants was developed, and the 5 suitable TPH variants were purified followed by immediate screen set-up eliminating the freezing-thawing step. No crystals of any

TPH2 variant were found. However, the pipeline screening produced leads for r*ch*TPH1 and crystals for c*ch*TPH1. The initial crystallization conditions identified for c*ch*TPH1 should be optimized to yield single crystals.

6 Mechanistic studies

In this chapter, the experiments that investigate the mechanism of TPH will be presented. Several different techniques can be used to investigate the mechanism of enzymes. Here, focus has been on the pre-reaction events, specifically substrate and cofactor binding including binding affinity and structural changes in the active site. Also, mutations have been made in the active site to investigate the importance of the mutated amino acids for the reaction mechanism. When studying mechanism and substrate binding, it is important to utilize several different techniques to get reliable results. Even though a crystal structure shows a specific feature, it might not represent the natural state of the free enzyme in solution. Hence, a combination of different techniques gives the most trustworthy results. Up until now, many of the available characterization methods could not be applied to TPH, due to the lack of pure and stable variants that can be concentrated to the high concentrations necessary. The development of the purification procedure for c*h*TPH2, which results in high yields of a stable and pure product that can be concentrated to 2.5 mM, opened up the possibility to utilize many of these techniques.

6.1 Mechanistic mutations

In this project, mutations of amino acids in the active site of c*h*TPH2 have been made in order to investigate their importance for the activity of the enzyme. c*h*TPH2 was chosen as the preferred variant to make mutations in, due to the high stability and yield from purification. To decide which mutations should be studied, a screening of the relevant literature was performed. Some mutational studies performed on TPH, PAH, and TH will be presented here. First, the disease related mutations of TPH will be discussed, follow by a brief review of mutations of TPH, PAH, and TH which gives insight into the mechanism of the AAAHs. Last, the background for the 4 mutants made in this project, will be described.

6.1.1 TPH disease related mutants

Both TPH1 and TPH2 are linked to a number of disorders in humans. Thus, polymorphisms in TPH related to these disorders have been widely investigated. In 2009, McKinney et al reviewed and characterized all reported missense variants for TPH2(*146*). Four of these were present in the catalytic domain: Pro206Ser (bipolar disorder), Arg303Trp (ADHD), Ala328Val, and Arg441His. None of these 4 mutations were close to iron in the active site, but they still had a significant effect on activity(*146*)(*102*). As the reported disease related mutants of TPH2 have all be studied by other groups, they were not chosen for mutation in this project.

6.1.2 Mutations for investigation of AAAH mechanism

Mutational studies have been used to shed light on the reaction mechanism of AAAHs. In the following a few examples of this will be given.

Mutations of several amino acids in the active site of rat TH have been performed. The mutant Glu332Ala catalyzed the uncoupled reaction, where only the cofactor and not the amino acid is hydroxylated(*70*). Kinetic studies showed that this amino acid has an important role in directing protonation of the Fe-OO-pterin bridge ensuring correct formation of the hydroxylating intermediate(*70*). Without Glu332 the hydroxylating intermediate is not formed, and tyrosine cannot be hydroxylated as in the coupled reaction. Another mutation study of rat TH on Glu332, finds that it greatly affects the binding affinity of the cofactor, and also is important for ensuring a coupled reaction(*147*). For PAH, the corresponding amino acid Glu286 has also been mutated, which confirmed its importance in catalytic activity and correct positioning of the cofactor(*130*). For this important glutamate residue, several different techniques were used to confirm the importance for the reaction mechanism by mutation.
Another example of mutational studies was performed on disease-related amino acids in the active site of PAH using spectroscopic techniques combined with kinetic analysis. The mutant PAH Glu280Lys remained 6C in iron coordination upon substrate and cofactor binding in contrast to wild-type PAH, which changes from 6C to 5C(*84*). Another mutant, Arg158Gln, did change iron coordination to 5C, but still the reaction catalyzed was mainly uncoupled. This indicates that the first step of the mechanism is formation of a peroxy-pterin species, which subsequently reacts with the iron site if the pterin is properly oriented for formation of a Fe-OO-pterin bridge and an open coordination position is available on the iron. Otherwise, the reaction becomes uncoupled, as no hydroxylating intermediate is formed.
As mentioned in section 2.3.2.1, mutational studies can also clarify the effect of different amino acids on substrate specificity(*62*).

6.1.3 Background for the mutants made in this project

Based on the screening of the literature, 2 mutations of each 2 amino acids in c*h*TPH2 were chosen for this project.
From the PAH crystal structures, Tyr325 was identified as a possible important amino acids, as it swings out with an iron-coordinated water when the geometry changes from 6C to 5C upon substrate and cofactor binding (see section 2.4.1, Figure 10). Earlier DFT studies have also confirmed the importance of this amino acid in the hydrogen bonding network in the active site of TPH(*76*). This tyrosine residue (Tyr358 in *h*TPH2 numbering) is placed in the second coordination sphere from iron and is conserved in all AAAHs except for human TH, where it is replaced by serine conserving the hydroxyl-group functionality. This tyrosine residue has been mutated to Phe in PAH without any significant loss in activity(*130*) thus excluding any direct functional role in catalysis. For rat TH the corresponding amino acid has also been mutated to Phe and the mutant showed full activity in the hydroxylation of tyrosine, leading to the conclusion that it had no essential mechanistic role(*148*). However, in 2003 mass spectrometric analysis of human Tyr325Phe PAH showed that the mutated phenylalanine had been post-translationally hydroxylated back to tyrosine eliminating the difference between the mutant and wild-type proteins. For bacterial Tyr325Phe PAH there was no such repair mechanism, and the specific activity was reduced 30-fold compared to the wild-type(*149*). These results suggest that Tyr325 might have an important role despite earlier results. This was investigated further in 2005, where Tyr325 in PAH was mutated to Leu, Ala, Ser and Phe(*150*). Only Tyr325Leu and Tyr325Phe mutations resulted in stable enzymes. Compared to wild-type PAH the Tyr325Leu showed decreased specific activity, decreased coupling efficiency and decreased iron content, suggesting that Tyr325 does indeed play an important role. The importance of this tyrosine residue for the

catalytic activity in TPH can be investigated experimentally by making mutations of this amino acid and testing the mutants for activity.

The c*h*TPH2 mutants Tyr358Phe and Tyr358Ser were chosen for the mutational study. The overall goal was to investigate whether the results from PAH are also valid for TPH. Tyr358Phe c*h*TPH2 was made in order to check for self-hydroxylation and effect on activity. Tyr358Ser c*h*TPH2 was made because this mutant has conserved the OH group, which might be important for functionality, but has removed the aromatic ring, which might have an effect on stability. Furthermore, the corresponding residue in human TH is a serine.

Mutational studies can also be used to investigate the plasticity of the iron coordinating amino acids in TPH. The 3 amino acids coordinating iron in the facial triad have been subject to mutations in both PAH and TH in earlier studies(*151–154*). Mutation of all three amino acids to glutamine, glutamate or histidine in PAH shows that the activity is most sensitive to mutations in the glutamate ligand, but that all mutations gave low activities and reduced affinity for iron binding(*151*). The iron ligands in TH are much less sensitive to mutations than the iron ligands in PAH(*151*). However, in both enzymes mutation of glutamate is more disruptive to catalysis than mutation of either of the two histidines.

Here, the glutamate iron ligand mutants Glu363Gln c*h*TPH2 and Glu363His c*h*TPH2 were chosen. The mutation to glutamine was thought to partly conserve the bidentate binding functionality of glutamate, which is important in the conformational changes in the active site upon substrate and cofactor binding. The amide of glutamine could coordinate the iron through both the deprotonated nitrogen and the oxygen, although this is not a very likely scenario. The mutation of glutamate to histidine is inspired by the existence of other oxygen-activating enzymes with a 3-His binding iron motif(*155*, *156*), and thus it was tested if TPH can hydroxylate tryptophan with a 3-His iron coordination.

6.1.4 Experimental procedures

Most of the experimental work was performed by bachelor student Anne Kathrine Nielsen and generally follows the same procedures as for wild-type c*h*TPH2.

6.1.4.1 Expression

The 4 c*h*TPH2 mutants Tyr358Phe, Tyr358Ser, Glu363Gln and Glu363His were expressed as described in section 4.2.1, following the procedure for wild-type c*h*TPH2.

6.1.4.2 Purification

The 4 c*h*TPH2 mutants were purified following the procedure and buffer systems for wild-type c*h*TPH2, see section 4.2.4.1, using an anion-exchange step followed by a gel filtration step. Only one tube of cells corresponding to 650 mL culture was purified, and thus the smaller columns Q Sepharose High Performance 16/10 and HiLoad Superdex 75 16/60 prep grade were used. In the anion-exchange a less steep gradient was used (0-15 % buffer B over 20 CV), to increase resolution. Furthermore, the protein sample was stored overnight at 4 °C after the anion-exchange, to allow for SDS-PAGE analysis of the peaks in the chromatogram.

Wild-type c*h*TPH2 was purified as described in section 4.2.4.1.

The protein concentration was determined using the theoretical extinction coefficients(*157*) seen in Table 20 below by measurement on a Nanodrop 1000 spectrophotometer from ThermoScientific.

Table 20: Theoretical extinction coefficients used for protein concentration determination of the four c*h*TPH2 mutants.

Protein	Extinction coefficient at 280 nm
Wild-type c*h*TPH2	39310 M^{-1} cm^{-1}
Tyr358Ser c*h*TPH2	37820 M^{-1} cm^{-1}
Tyr358Phe c*h*TPH2	37820 M^{-1} cm^{-1}
Glu363His c*h*TPH2	39310 M^{-1} cm^{-1}
Glu363Gln c*h*TPH2	39310 M^{-1} cm^{-1}

6.1.4.3 SDS-PAGE

See section 4.2.5.1.

6.1.4.4 Mass spectrometric analysis

Using Micro Bio-Spin 6 Chromatography columns from Bio-Rad, the purified protein was transferred to a 0.6 M NH_4OAc solution (diluted from a prefabricated 7.5 M NH_4OAc solution in plastic bottle from Sigma Aldrich). For denatured samples, the protein was transferred to a 2 (v/v)% formic acid solution. The samples were filtered through a 0.22 µm Millex-GV syringe filter. The mass spectrometer used was a LCT premier instrument with a nano-ESI source. The following instrument settings were used: Capillary voltage 1100-1600 V and ionguide voltage 70-150 V. The instrument was calibrated using 100 mg/mL cesium iodide in 50 (v/v)% isopropanol. The function MaxEnt1 in the software suite MassLynx 4.1 was used to determine the mass of the detected ions.

6.1.4.5 Activity measurements

The activity measurements were performed as described in section 4.2.5.2. Protein concentrations between 0.5 and 3.3 µM were used. For blank measurements, 20 mM HEPES/NaOH, 100 mM $(NH_4)_2SO_4$, pH 7.2 was used, and this buffer was also used for dilution of protein samples.

6.1.5 Results and discussion

In the following section the results from the experimental work with the 4 c*h*TPH2 mutants will be presented and discussed.

6.1.5.1 Solubility test

In order to compare the expression levels and solubility between the different mutants, an SDS-PAGE solubility test was made loading raw extract, resuspended pellet and supernatant. The result is seen in Figure 85. The expression levels of the 4 mutants look similar based on the intensity of the band around 36 kDa in the raw extract samples. As the samples were run on different gels with different staining/destaining solutions and times, the comparison can only be approx. A large difference in solubility is however seen between the mutants. The raw extract and pellet samples can be directly compared, as pellet was resuspended in an equal volume of buffer as the raw extract. Clearly, Tyr358Phe c*h*TPH2 is the most soluble mutant, with only a small fraction of the protein in the pellet. Tyr358Ser c*h*TPH2 is on the other hand not very soluble. This corresponds well with the fact that Tyr358 is placed in the hydrophobic interior of the protein, and mutation to a hydrophobic amino acid such as Phe does not alter the solubility as much as the mutation to Ser, which is a polar/hydrophilic amino acid. The introduction of a hydrophilic amino acid in the

interior of a protein can cause folding problems and change the overall structure. Tyrosine is itself defined as a polar amino acid due to the hydroxy-group, but the aromatic ring makes it partly hydrophobic. Another explanation for the difference in solubility could be that the aromatic ring structure found in both Tyr and Phe, but not in Ser, is important for the stability of the protein. Aromatic rings can participate in stacking interactions with other aromatic ring amino acids (see Figure 91), and can through this contribute to the structural stability of the protein. From the crystal structure of c*h*TPH1 with BH_2 bound (PDB ID: 1MLW) no clear π-stackings interactions are seen for the corresponding tyrosine. Either way, it can be concluded that Tyr325 is important for solubility of TPH. This is in agreement with the mutation studies of PAH from 2005, where mutation of the corresponding tyrosine to serine resulted in enzyme that aggregated quickly after purification, while mutation to Phe gave stable protein with similar yield as for wild-type PAH(*150*).

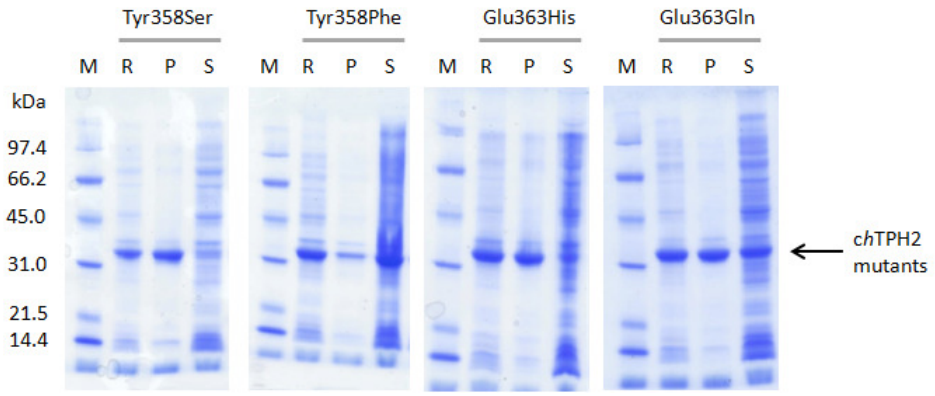


Figure 85: SDS-PAGE analysis of the solubility test of the four mutants of c*h*TPH2. M: Molecular weight standard, R: Raw extract, P: Pellet, S: Supernatant. For Tyr358Ser, Glu363His, and Glu363Gln, 0.5 µL R and P and 3 µL S was loaded, while for Tyr358Phe, 1 µL R and P and 5 µL S was loaded.

Glu363His c*h*TPH2 and Glu363Gln c*h*TPH2 are both not very soluble, with a large fraction of the protein in the pellet. However, there are clear bands of protein in the supernatant, which allows for purification. Since Glu363 is placed at the active site and is involved in iron binding, it is not surprising that this amino acid also has a great effect on solubility of the protein. The iron binding amino acids are conserved among all AAAHs, suggesting the importance of these. For PAH, mutation of the corresponding glutamate to histidine or glutamine also showed that a significant amount of protein was found in insoluble inclusion bodies(*151*). For wild-type c*h*TPH2 approx. 90 % of the protein is found in the soluble fraction(*13*).

6.1.5.2 Purification

The 4 c*h*TPH2 mutants were purified following the procedure for purification of c*h*TPH2 consisting of an anion-exchange step followed by a gel-filtration step. The chromatograms from the anion-exchange can be seen in Figure 86. For all mutants, SDS-PAGE analysis was made of all peaks in the chromatogram in order to decide which peak(s) contained TPH. Only the SDS-PAGE for the purification of Tyr358Ser c*h*TPH2 is shown in the figure.

Due to the difference in solubility between the mutants, the peak height also differs in the anion-exchange. However, all chromatograms have the same overall features, and all 4 c*h*THP2 mutants elutes at conductivities around 10 mS/cm. Another common feature is the poor separation of peaks which is mainly due to the low concentration of target protein compared to impurities.

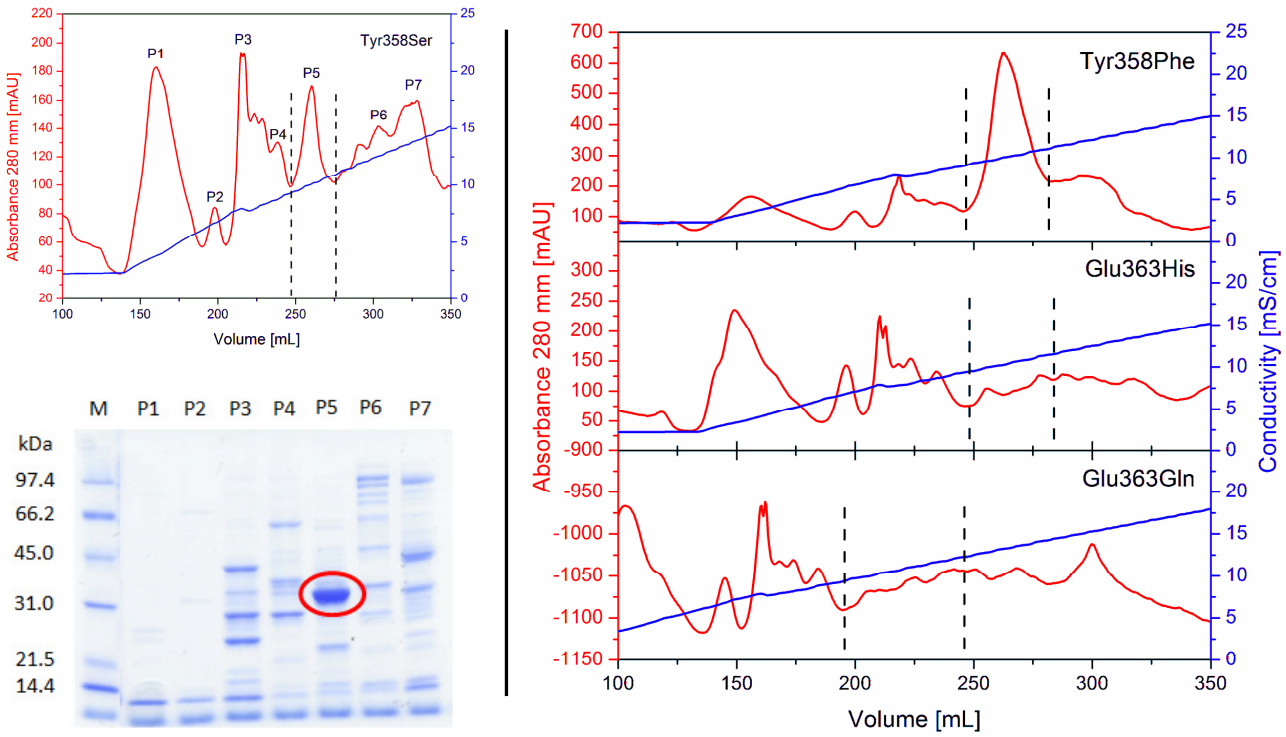


Figure 86: Anion-exchange of the four chTPH2 mutants. Left: Chromatogram and SDS-PAGE analysis from the anion-exchange of Tyr358Ser chTPH2 on a Q Sepharose High Performance 16/10 column. M: Molecular weight standard, P1-P7: Samples from the corresponding peaks in the chromatogram. Right: Chromatograms for the anion-exchange of the three other chTPH2 mutants. The fractions marked with dashed lines contain TPH and are collected for further purification. The negative values on the absorption axis in the chromatogram of Glu363Gln are due to an error in zero setting.

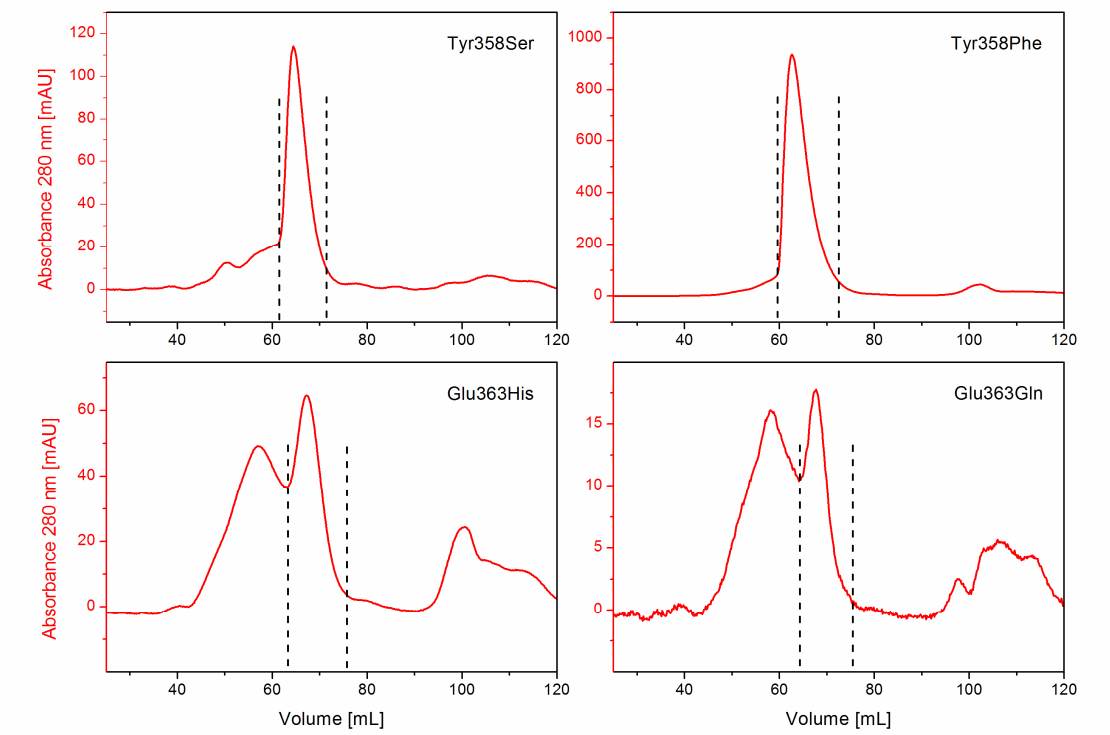


Figure 87: Chromatograms from the gel-filtration of the four *ch*TPH2 mutants on a HiLoad Superdex 75 16/60 prep grade column. The fractions marked with dashed lines contain TPH and are collected as the final product. Notice the difference in peak intensity between the four mutants.

In the gel-filtration, all mutants elute around 65-70 mL (see Figure 87) corresponding to a mass of approx. 40 kDa calculated from calibration curves. This corresponds well with the molecular weight of c*h*TPH2 (approx. 36 kDa), suggesting that the mutants exists as monomers as expected.
The yield of purifications of the four c*h*TPH2 mutants can be seen in Table 21, where the yield of wild-type c*h*TPH2 is included for comparison. It is clear that the yields of all mutants are significantly lower than for the wild-type, probably due to lower solubility. The most soluble mutant, Tyr358Phe, also gives the highest yield. If more protein is needed for further characterization, 3 tubes of cells should be purified at one time using larger columns.

Table 21: Yields of the purifications of the four c*h*TPH2 mutants. Concentration is defined as that of the final product after concentration with ultrafiltration.

Protein	Yield [mg/L culture]	Concentration [µM]
Wild-type c*h*TPH2	70(*13*)	2500
Tyr358Ser c*h*TPH2	2	14
Tyr358Phe c*h*TPH2	36	90
Glu363His c*h*TPH2	3	19
Glu363Gln c*h*TPH2	1	7

6.1.5.3 Characterization of purification products

To check the purity of the purification products of the four c*h*TPH2 mutants, an SDS-PAGE analysis was conducted. The result is seen in Figure 88. Based on the SDS-PAGE all 4 mutant solutions appear pure, but there are traces of some impurity in the Glu363His c*h*TPH2 sample with molecular weight of approx. 40 kDa. However, the samples are pure enough for kinetic analysis.

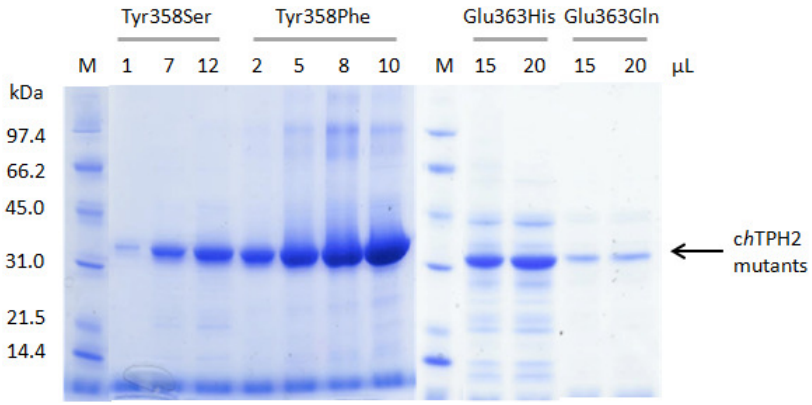


Figure 88: SDS-PAGE analysis of the final products from the purifications of the four c*h*TPH2 mutants. M: Molecular weight standard. Numbers: µL loaded.

It was attempted to analyze all four mutant solutions with mass spectrometry. This did not succeed for Tyr358Ser c*h*TPH2, Glu363His c*h*TPH2, and Glu363Gln c*h*TPH2 probably due to the very low concentrations of the purified products, see Table 21 . During exchange into NH_4OAc buffer the sample was diluted, and attempts to re-concentrate the samples did not succeed. A concentration of approx. 5-10 µM is favorable for mass spectrometry.

For Tyr358Phe c*h*TPH2, mass spectrums of both the denatured and intact protein were made. The spectrum of the denatured protein can be seen in Figure 89. The mass of denatured Tyr358Phe c*h*TPH2 was calculated to 36300.3 Da, which corresponds very well with the theoretical mass of apo-Tyr358Phe c*h*TPH2 of 36300.5 Da. If the mutated Phe had been hydroxylated back to Tyr, the theoretical mass would be 36316.5 Da. From the mass spectrometric measurement it can thus be concluded that Phe358 is not post-translationally hydroxylated back to Tyr in c*h*TPH2.

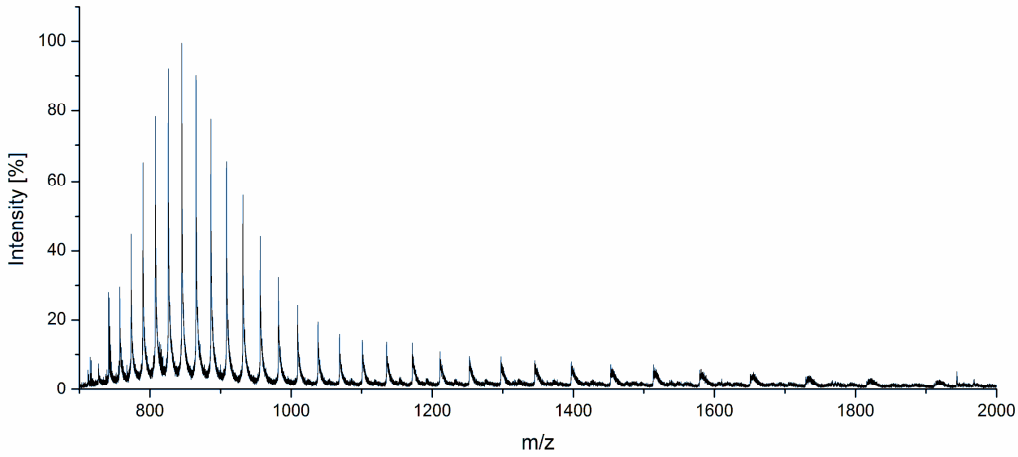


Figure 89: Mass spectrum of 10.6 µM denatured Tyr358Phe c*h*TPH2 in 2 (v/v)% formic acid. The calculated mass is 36300.3 Da.

The spectrum of intact Tyr358Phe c*h*TPH2 can be seen in Figure 90. The mass of the major species was calculated to 36362.7 Da. This is in agreement with the theoretical mass of holo-Tyr358Phe c*h*TPH2 of 36353.1 Da. The smaller peak left of the main peak (visible in the inset of Figure 90) represents apo-Tyr358Phe c*h*TPH2, with a calculated mass of 36300.0 Da. Thus, the protein is not fully loaded with iron, or some protein loses iron during the ionization process.

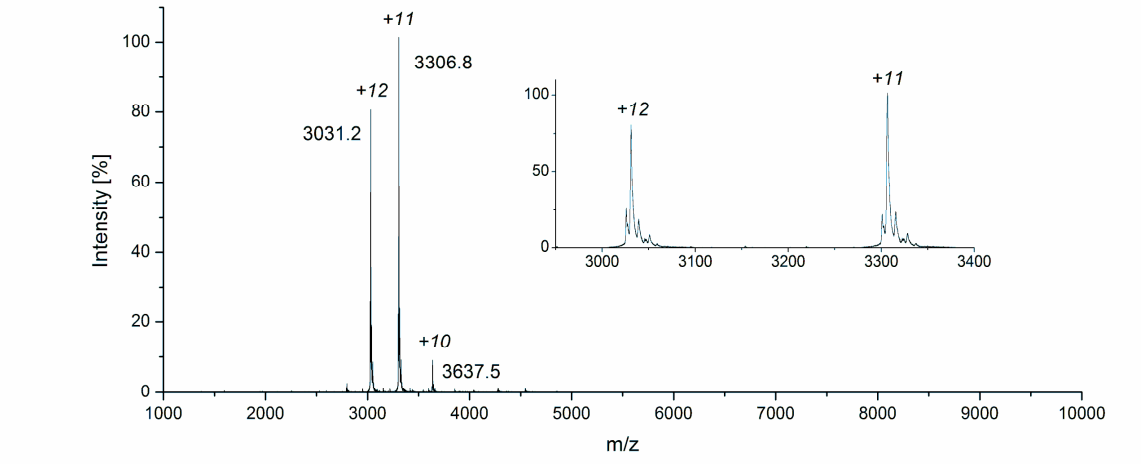


Figure 90: Mass spectrum of 6.9 µM Tyr358Phe c*h*TPH2 in 0.6 M NH_4OAc. The calculated mass is 36362.7 Da.

6.1.5.4 Activity measurements

The purified samples of the four c*h*TPH2 mutants were tested for tryptophan hydroxylation activity and compared to wild-type c*h*TPH2. The results can be seen in Table 22.

Table 22: Results from the activity measurements of the four c*h*TPH2 mutants. The specific activity is calculated by dividing the activity with the protein concentration of the sample.

Sample	Activity [Int/min]	Concentration [µM]	Specific activity ± st. dev. [Int/min/µM]
Blank	120	-	-
Wild-type c*h*TPH2	790 590 650 490	0.5 0.5 0.5 0.5	1300 ± 300
Tyr358Ser c*h*TPH2	100 220 240 260	0.5 2.5 3.0 3.0	110 ± 50
Tyr358Phe c*h*TPH2	510 410 310 300	0.5 0.5 0.5 0.5	800 ± 200
Glu363His c*h*TPH2	20 110	0.5 3.3	40 ± 5
Glu363Gln c*h*TPH2	30 60	0.5 1.7	50 ± 20

From the activity measurements, it can be concluded that both Glu363 c*h*TPH2 mutants are inactive in the hydroxylation of tryptophan, as the measured activities are not significantly higher than the blank. The nature of the assay and the low protein concentrations set limitations to how low activities that can be measured. Thus, it is possible that the two mutants have very low activities outside the range of this assay. Mutation of the corresponding glutamate in PAH resulted in very low affinity for iron with K_{Fe} in the mM range(*151*). In the assay used here, 25 µM iron is added to the assay. Assuming that iron affinity is the same as for PAH and using the protein concentration from the assay, only a very small fraction of the mutant proteins would bind iron. Thus, the lack of activity in the assay may be due to low affinity for iron. In future experiments, this should be tested using a large excess of iron in the assay. It can be concluded that glutamate in the 363 position of c*h*TPH2 is essential for activity and/or iron affinity, and that c*h*TPH2 cannot hydroxylate tryptophan with a 3-His triad.

Both Tyr358Ser c*h*TPH2 and Tyr358Phe c*h*TPH2 show significant activities in the assay, but are still less active than the wild-type. Thus, Tyr358 is important, but not essential, for activity in c*h*TPH2. Tyr358Phe is much more active than Tyr358Ser and has activity in the same range as wild-type, suggesting that the aromatic ring conserved in both Tyr and Phe is important for optimal activity. Earlier studies of PAH have shown that the corresponding tyrosine is involved in cofactor binding and positioning through hydrophobic interactions(*130*). It is also speculated that this tyrosine through the hydrogen bond to an iron coordinated water (see Figure 91) can affect the catalytic activity and stability of the iron active site(*150*). Based on the kinetic experiments in this project, these two effects cannot be distinguished. However, as the mutation that conserves the aromatic ring results in higher activity that the mutation that conserved the hydroxyl group, it can be speculated that the hydrophobic interactions between Tyr358 and the cofactor are more important than the hydrogen bond to the iron coordinated water.

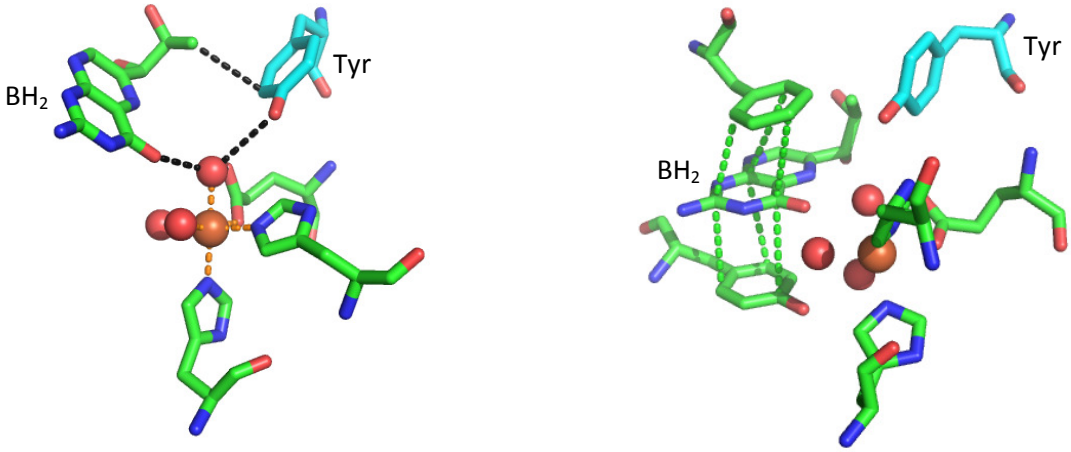


Figure 91: The active site of c*h*TPH1 with BH_2 bound (PDB ID: 1MLW). Iron is shown as an orange sphere and water as red spheres. Tyrosine corresponding to Tyr358 is shown in cyan. Left: Hydrophobic interaction and hydrogen bonding between Tyr358, BH_2 and iron coordinated water. Right: Example of π-stacking interaction between Phe241, BH_2 and Tyr235 (*h*TPH1 numbering). The figure was made using PyMOL(*50*).

6.1.6 Conclusions

Four mutants (Tyr358Ser, Tyr358Phe, Glu363His, Glu363Gln) of c*h*TPH2 have been successfully expressed and purified using procedures for wild-type c*h*TPH2. The mutants are all less soluble than wild-type, with Tyr358Phe c*h*TPH2 as the most soluble. Mutation of the iron coordinating Glu363 to histidine or glutamine gave inactive protein, and shows that glutamate is essential for catalytic activity. Whether the inactivity is caused by a lower affinity for iron or catalytic dysfunction cannot be concluded. Mutations of Tyr358 to serine or phenylalanine yielded active protein, with Tyr358Phe as the most active. Thus, it can be concluded that Tyr358 is important, but not essential, for catalytic activity. The Tyr358Phe c*h*TPH2 mutant was not hydroxylated back to tyrosine during expression and purification, as is seen for human PAH(*149*).

6.2 Circular dichroism and magnetic circular dichroism

As described in the introduction to TPH in chapter 2, both PAH and TH change coordination of iron from 6 coordinated (6C) to 5 coordinated (5C) when both substrate and cofactor bind. Such investigations have until now not been possible for TPH, due to the lack of a stable variant that can be concentrated to the required mM concentrations. With the development of the purification procedure for c*h*TPH2, the possibility of such investigations suddenly arose.

Thus, in order to investigate the coordination of iron in the active site of TPH upon binding of substrate and/or cofactor, circular dichroism (CD) and magnetic circular dirchroism (MCD) experiments were performed on c*h*TPH2. The experiments were made in collaboration with graduate student Lei V. Liu, The Solomon Laboratory, Stanford University, Palo Alto, California, USA under supervision of Professor Edward I. Solomon.

6.2.1 Introduction to the techniques

The utilization of CD and MCD to investigate Fe(II) in the active site of proteins is relatively new, thus an introduction to the techniques will be given. The introduction is based on reference (*158–162*).

6.2.1.1 Electronic configuration of Fe(II) in biological systems

Ferrous iron, which is the catalytically active species in the AAAHs, has 6 d electrons. Fe(II) in the AAAHs is paramagnetic high spin, due to a low ligand field (LF) splitting compared to electron pairing energy caused by weak nitrogen and oxygen ligands(*158*). The energies of the 5 d orbitals depend on the geometry of the iron complex. The Fe(II) free ion has a 5 d orbital ground state which in an octahedral field splits into a triply degenerate ($^5T_{2g}$) ground state and a doubly degenerate (^5E_g) excited state. Due to the spin selection rule (the excited electron cannot change spin in a transition) only the one electron with opposite/down spin in the ground state can be excited, since all excited state d-orbitals are already filled with spin-up electrons, see Figure 92. Thus, LF transitions from one ground state to either of the two excited states may occur. In the low-symmetry environment of non-heme iron proteins active site, the $^5T_{2g}$ and ^5E_g will split further in energy dependent on the coordination number and geometry of the Fe(II) ligation sphere(*162*) as shown in Figure 92. For a distorted 6C site the ^5E_g state will split by ~2000 cm^{-1} resulting in two ligand field transitions from the $^5T_{2g}$ state to each of the two components of the ^5E_g excited state. For the biologically relevant oxygen and nitrogen ligands of non-heme Fe(II) sites, the transitions will be centered around 10000 cm^{-1}. In a 5 C square pyramidal site the ^5E_g state will split further resulting in two LF transitions at ~10000 and ~5000 cm^{-1}, while distortion of the 5C site to trigonal pyramidal changes the LF and lowers the energies of the two LF transitions to <10000 and <5000 cm^{-1}. A distorted tetrahedral 4C site will result in two low energy LF transitions around 4000-7000 cm^{-1}.

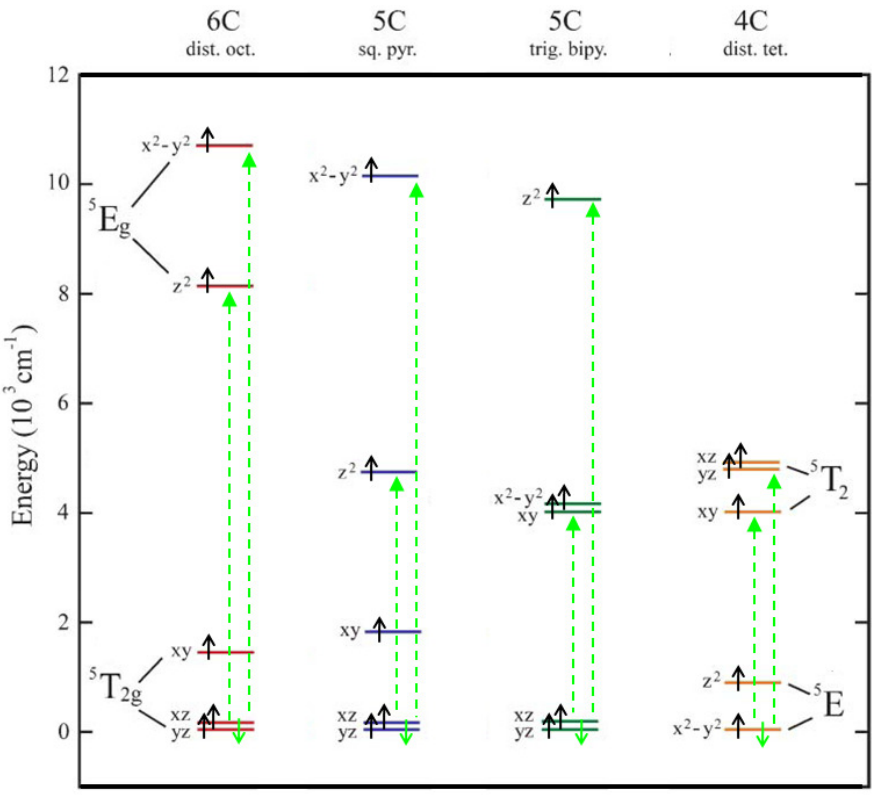


Figure 92: The electronic transitions dependent of the coordination geometry of Fe(II). Adapted with permission from (*161*)**. Copyright 2000 American Chemical Society.**

By determining the energy of the electron d-d (LF) transitions, the splitting of the exited state d-orbitals can be determined. This gives insight into the geometric structure as the splitting depends on the coordination geometry of iron according to ligand field theory.
However, the splitting of the excited state d-orbitals can be hard to determine experimentally. The d-d transitions are difficult to observe in absorption due to the parity forbidden nature, low extinction coefficients and falling in the near infra-red (near-IR) region were they are disturbed by intense contributions from protein and buffer vibrations(*160*). Recently, experimental methods to probe these transitions have been developed using CD and MCD(*158*).

6.2.1.2 Circular dichroism and magnetic circular dichroism spectroscopy

Circular dichroism (CD) and magnetic circular dichroism (MCD) are techniques that can be used for studying electronic transitions in metal complexes. In this project near-IR CD and MCD have been used to study the electronic and geometric structure of the active site iron in the catalytic domain of c*h*TPH2 by observing the d-d electron transitions described in the previous section.
In both CD and MCD spectroscopy the difference in absorbance of left-handed versus right-handed circularly polarized light is measured. The difference is recorded as a function of wavelength, which in the near-IR case goes from 600-2000 nm. In MCD the sample is placed in a strong magnetic field parallel to the propagation direction of the circular polarized light.
For a material to exhibit CD activity is must possess chirality(*158*), which means that the iron coordination moiety cannot be superimposed on its mirror image. Optical activity can also be induced by applying a magnetic field, so that all substances exhibit MCD activity(*158*). CD and MCD probe the same electronic transitions and the results from the two techniques should be comparable.

6.2.1.2.1 Linearly and circularly polarized light

Linearly or plane polarized light can be obtained as a superposition of a left circularly polarized (LCP) wave and a right circularly polarized (RCP) wave with identical amplitudes. For linearly polarized light the magnitude of the electric field vector oscillates only in one plane. For circularly polarized light the electric field vector has a constant length, but rotates around its propagation direction creating a helix or spring shaped wave. Circularly polarized light is defined as right-handed (or simply right) when the electric field vector rotates clockwise as viewed by an observer facing the direction of the light propagation. The magnetic field vector is always perpendicular to the electric field vector. Both the electric and magnetic fields oscillate in a plane perpendicular to the direction of propagation.
CD and MCD effects are observed when a substance has a different absorbance for RCP in comparison to LCP at a certain wavelength. The plane polarized light that is shone upon the sample will be affected by this and the outgoing light will be an elliptical wave instead of a plane polarized wave, see Figure 93. This phenomenon arises because the intensity and thereby the amplitude of the two circularly polarized waves no longer are identical, due to the difference in absorption.

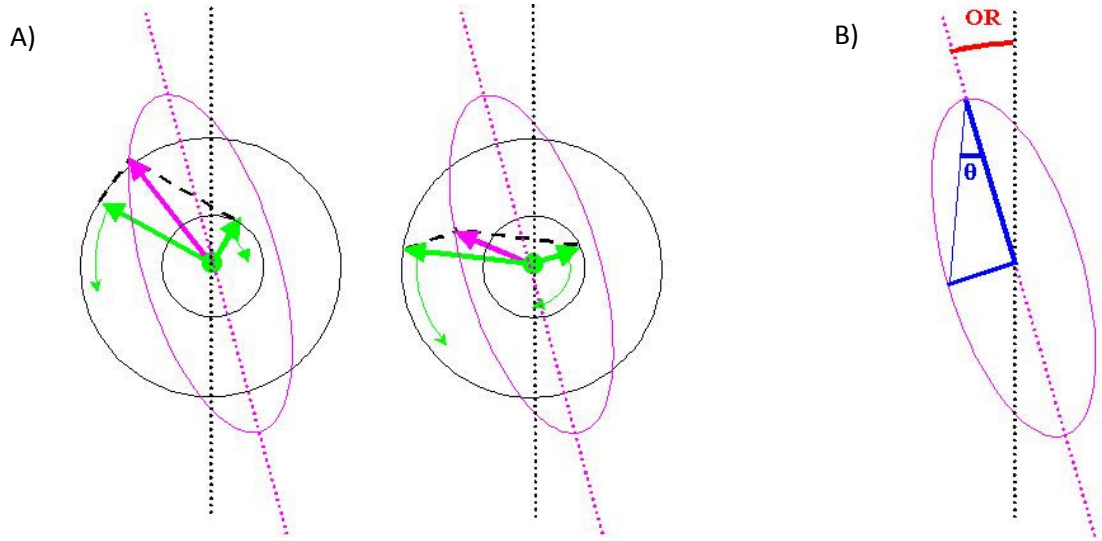


Figure 93: A) The elliptical wave (pink) resulting from difference in amplitude of the two circularly polarized light vectors (green). B) Definition of ellipticity (θ) and optical rotation (OR) of the elliptical wave. Adapted from (*163*)**.**

The elliptic trajectory of the resulting wave can be described by θ called ellipticity and OR called optical rotation. As seen from Figure 93, θ gives the ratio of the minor to the major axis of the ellipse by the angle of the tangent. OR, sometimes referred to as α, is the angle between the initial plane of polarization and the major axis of the ellipse of the resultant light.

The value that is usually measured in a CD/MCD experiment is *ΔA*, the difference in absorbance of LCP and RCP light.

$$\Delta A = A_L - A_R$$

Lambert-beers law can be applied to give:

$$\Delta A = (\varepsilon_L - \varepsilon_R) \cdot c \cdot l$$

where ε_L and ε_R are the molar extinction coefficients for LCP and RCP, respectively, c is the molar concentration and l is the path length in cm. The molar circular dichroism $\Delta\varepsilon$ is given as:

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R$$

Though *ΔA* is often the measured quantity, most CD/MCD results are reported in degrees of ellipticity for historical reasons. Molar ellipticity [θ] is proportional to molar CD Δε by the following equation:

$$[\theta] = 3298.2 \cdot \Delta\varepsilon$$

This relationship is derived from the definition of the ellipticity of the polarization(*164*) and the unit of Δε is M^{-1} cm^{-1}.

6.2.1.2.2 Circular dichroism

In CD, the absorption of RCP and LCP light is caused by transitions of electrons from the ground state to the excited states. In order for an electronic transition to give rise to optical activity as measured by CD spectroscopy, the transition must be both electric and magnetic dipole allowed according to the selection rules(*165*).

The first selection rule is the spin selection rule ΔS = 0 for spin-allowed transitions. This means that the excited electron cannot change spin in a transition. For high-spin Fe(II) d^6 in an octahedral ligand field, this means that only the one electron with opposite spin in the ground state can be excited, since all excited state d-orbitals are already filled with electrons. This one transition is thus magnetic dipole allowed. However, it is electric dipole forbidden according to the second selection rule.

The second selection rule says that all d-d transitions are parity forbidden according to the LaPorte selection rule, where transitions within the same subshell are forbidden in a centrosymmetric environment. However, metal sites in proteins usually have low symmetry with no inversion, which removes the parity restriction and splits the ^5E_g excited state into two non-degenerate states that can be probed by CD(*158*). For free iron ions in solutions, the environment is centrosymmetric (6 octahedrally coordinated waters) and thus the d-d transitions in these complexes are parity forbidden and will not be observed with CD.

6.2.1.2.3 Magnetic circular dichroism

In MCD, the difference in absorption of LCP and RCP light is also measured, but the sample is placed in a strong magnetic field (up to 7 T) parallel to the propagation direction of the circularly polarized light.
The MCD intensity is dependent on three different contributions, called the $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$-terms. $\mathcal{A}$ is any feature with a derivative bandshape caused by degenerate ground or excited states. $\mathcal{B}$ is any feature with absorption bandshape and temperature-independent intensity caused by intermediates states (mixing). And $\mathcal{C}$ is any feature with absorption bandshape that increases in intensity with decreasing temperature(*158*). In the experiments performed here we focus on the $\mathcal{C}$-term.

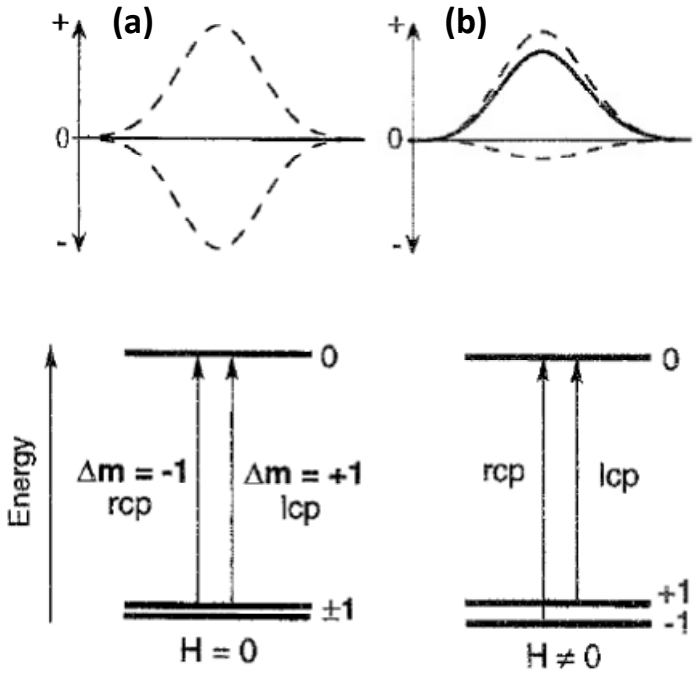


Figure 94. MCD $\mathcal{C}$-term mechanism. (a) With no applied field, the RCP and LCP transitions from a doubly degenerate ground state to an excited state cancel, resulting in no observed MCD intensity. (b) When a magnetic field is applied, the degeneracy of the two ground states is lifted. At low temperatures the ground states become unequally populated corresponding to the Boltzmann distribution. This results in an MCD transition with an absorption bandshape, called a $\mathcal{C}$-term. Adapted with permission from (*158*). Copyright 1995 Elsevier.

The $\mathcal{C}$-term arises from a degenerate ground state which is split in the magnetic field due to the Zeeman effect. Only paramagnetic compounds exhibit $\mathcal{C}$-term signals. As shown in Figure 94 LCP and RCP light induce different electronic transitions. With no magnetic field, the ground states are degenerate and the RCP and LCP induced transitions cancel out resulting in no observed MCD intensity (see Figure 94, a). In the presence of a magnetic field the degeneracy of the ground state is lifted. At low temperature the two levels of the ground state doublet are unequally populated corresponding to the Boltzmann distribution. Hence,

the intensities of the RCP and LCP transitions do not cancel anymore, leading to an absorption band shape for the $\mathcal{C}$-term. A further decrease in temperature or an increase in magnetic field results in a further increase in the population of the lowest-energy sublevel and therefore, the $\mathcal{C}$-term MCD intensity also increases.

All MCD intensity should increase linearly with magnetic field. $\mathcal{A}$- and $\mathcal{B}$-terms are temperature independent, while the $\mathcal{C}$-term is highly dependent on the temperature. Because the $\mathcal{C}$-term intensity is proportional to 1/T it is greatly enhanced at low temperatures compared to $\mathcal{A}$- and $\mathcal{B}$-term signals and dominates for paramagnetic transition metal complexes(*159*). $\mathcal{A}$-terms can generally be excluded as a source of MCD intensity for high-spin ferrous centers(*158*). The field and temperature dependence of MCD features should be checked during experiments to differentiate between $\mathcal{B}$- and $\mathcal{C}$-terms.

6.2.1.3 Model complexes

The $\mathcal{C}$-term feature will be observed for both transitions to the two excited state orbitals of AAAH Fe(II). This results in 2 positive peaks for a 6C octahedral Fe(II) complex within the near-IR region. In Figure 95 experimental MCD spectra are shown for different iron complexes with 6, 5 or 4 ligands.

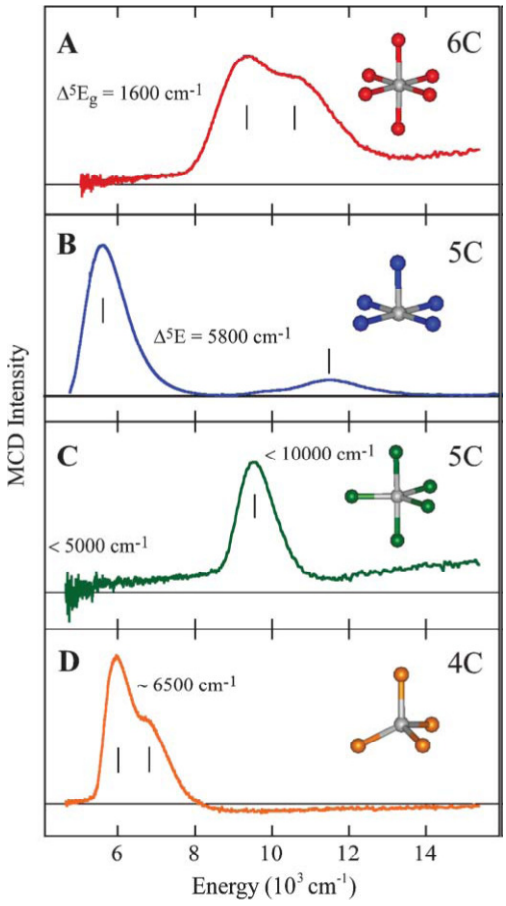


Figure 95: Low temperature MCD spectra of structurally defined Fe(II) model complexes. A: 6C octahedral $[Fe(H_2O)_6]$. B: 5C square pyramidal $[Fe(HB(3,5\text{-}{}_iPr_2pz)_3)(OAc)]$. C: 5C trigonal bipyramidal $[Fe(tris(2\text{-}(dimethylamino)ethyl)amine)Br]^+$. D: 4C tetrahedral $[Fe(HB(3,5\text{-}{}_iPr_2pz)_3)(Cl)$. Adapted with permission from (*161*). Copyright 2000 American Chemical Society.

From the experimental MCD spectra in Figure 95, the changes in LF transitions with changing geometry are clearly seen. Thus, the method can be used to determine the coordination of Fe(II) in non-heme iron enzymes such as the AAAHs.

6.2.1.4 Experimental considerations

The combination of CD and MCD has several advantages over e.g. absorption spectroscopy. Because of the parity-forbidden nature of the d-d transitions they exhibit low extinction coefficients and can therefore be hard to observe in absorption spectroscopy. However, CD and MCD spectroscopy are governed by different selection rules than absorption, which greatly favor the ligand field transitions. For CD, the magnetic dipole character of the d->d transitions of Fe(II) greatly enhances their CD intensity in the near-IR region relative to that of buffer and protein vibrations. Also, transitions in the CD spectrum have a sign as well as a magnitude which makes it easier to resolve overlapping bands contributing to an absorption feature(*159*).

The relevant transitions for Fe(II) complexes fall in the near-IR region where they are complicated by intense contributions from protein, H_2O and buffer vibrations in the <12000 cm^{-1} region. By exchanging protons with deuterium the vibrational overtone bands of O-H and C-H stretches are eliminated, which extends the transmission window from around 7000 cm^{-1} to 5300 cm^{-1}. Otherwise, the vibrational overtone bands of O-H and C-H stretches dominate the transmission spectra and can block light transmission throughout much of the near-IR region.

For MCD, high-spin paramagnetic Fe(II) will exhibit a $\mathscr{C}$-term MCD signal which at low temperature is 2-3 orders of magnitude more intense than the MCD signals associated with the diamagnetic protein background.

It is a requirement that the samples for CD and MCD must transmit light and must not depolarize the incident light. When measuring MCD at low temperatures (5K), the icing of aqueous protein samples will result in both loss of light transmission and depolarization. Therefore, a glassing agent, such as glycerol or sucrose, must be used to prevent icing. It should be checked that the glassing agent does not chemically alter the protein by comparing room temperature CD spectra with and without glassing agent.

As CD and MCD probe the same transitions, the peak positions should be comparable. However, due to the low temperature in MCD the peaks should be narrower in MCD than in CD, which is recorded at 5 °C.

As free Fe(II) ions in solution are octahedrally coordinated to six water molecules, they are not chiral and can therefore not be observed in CD spectroscopy. Thus, only features from the distorted active site iron will be visible in the spectrum, and there is no need to worry about excess iron. In MCD, however, features from all iron ions will be probed, and thus it is very important not to have free Fe(II) ions in the protein solution, as they will mix with the Fe(II)-active site signal, and make the results hard to interpret.

6.2.2 CD and MCD studies on AAAHS

Both PAH and TH have been studied by CD and MCD previously. The most important results from these studies will be presented here, along with the motivation for why a study of the geometric changes in TPH upon substrate binding is interesting.

The first AAAH to be investigated with CD and MCD was PAH in 1997(*69*). Solomon et al. studied PAH with and without substrate bound, and found that Fe(II) in PAH is 6C distorted octahedral. Binding of L-Phe in the active site resulted in geometric and electronic structural change at the iron center. The comparison of the PAH-resting and PAH-L-Phe spectra show that both are consistent with a distorted octahedral coordination. However, the transitions shift to higher energy for the substrate bound PAH corresponding to

a stronger ligand field. The presence of substrate in the active site pocket thus induces a stronger ligand field. The difference in energy between the two excited states decreased upon substrate binding.
In 1999, the interaction between the active site of PAH and the pterin cofactor was investigated again using CD and MCD(*67*). The addition of cofactor analogue 5-deaza-6-MPH$_4$ did not perturb the active site in the absence of substrate. However, in the presence of substrate the active site changed from 6C to 5C as seen by an intense low energy (~5000 cm^{-1}) peak indicative of a square pyramidal 5C iron site (see Figure 96, D). This was strong evidence that only when both substrate and cofactor have bound, an open coordination position on the iron is formed for dioxygen to react.
The 4 MCD spectra for the four PAH substrate/cofactor combinations can be seen in Figure 96 , A-D.

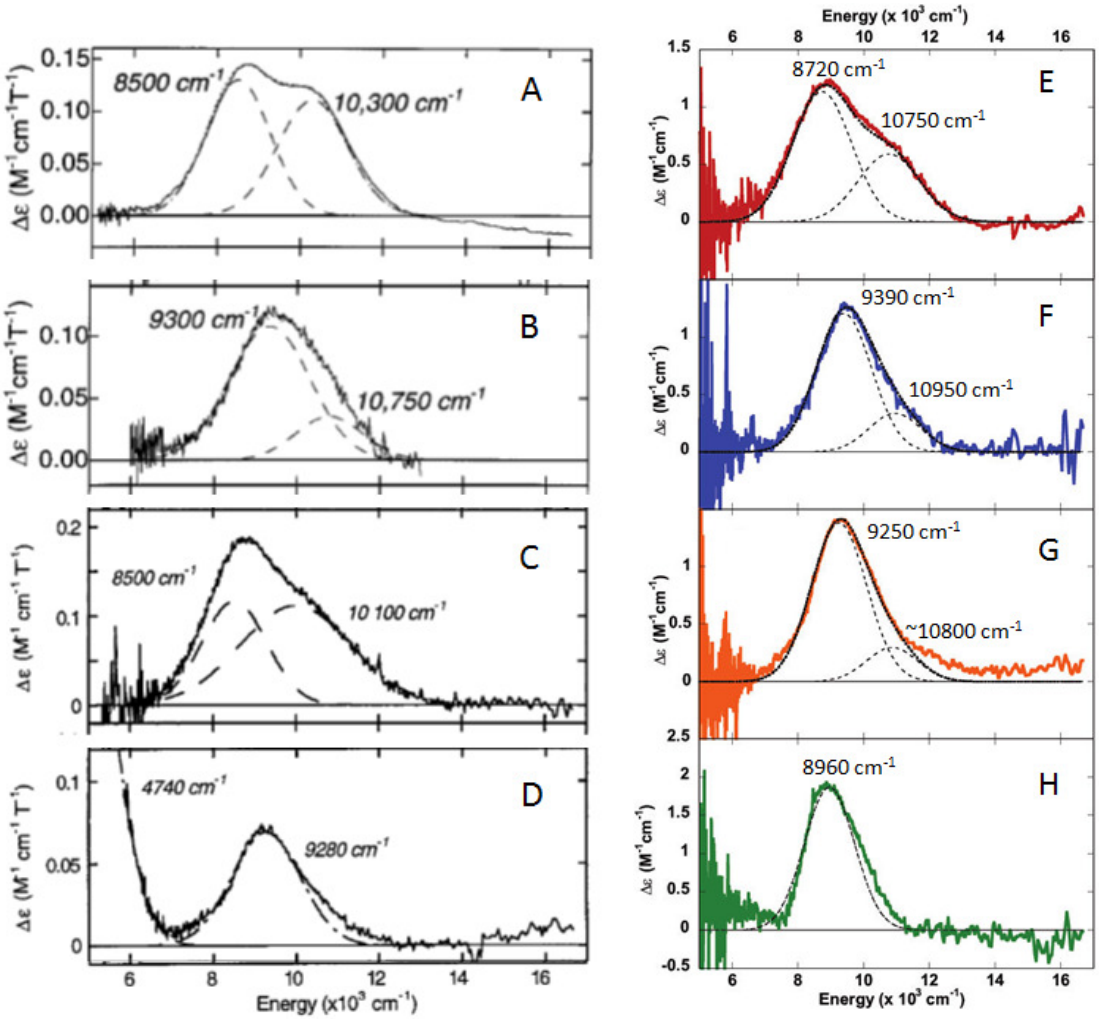


Figure 96: MCD spectra of the 4 substrate/cofactor combinations of PAH and TH. All spectra are measured at 5 K and 7 T. A-D: MCD spectra of the 4 PAH combinations. A: PAH. B: PAH + L-phenylalanine. C: PAH + 5-deaza-6-MPH$_4$. D: PAH +L-phenylalanine + 5-deaza-6MPH$_4$. Adapted with permission from(*69*) **and** (*67*)**. Copyright 1997+1999 American Chemical Society. E-H: MCD spectra of the 4 TH combinations. E: TH. F: TH + L-tyrosine. G: TH + 6MPH$_4$. H: PAH + L-tyrosine + 6MPH$_4$. Adapted with permission from** (*70*)**. Copyright 2009 American Chemical Society**

Later, in 2003, two phenylketonuria inducing mutants were investigated, also using MCD and XAS(*84*). This study showed that one of the mutants did not make the coordination sphere change of iron from 6C to 5C upon binding of substrate and cofactor analogue, and this could explain the lacking hydroxylation activity of the amino acid seen in kinetic analysis of this mutant.

Wild-type TH and two mutants have also been investigated using the above mentioned techniques(*70*). The results were very similar to the PAH results and can be seen in Figure 96, E-H. For the substrate only (F) and cofactor only (G) bound TH combinations, the MCD spectra showed distorted 6C geometry with slightly stronger ligand fields and less axial distortion compared to TH-resting sample, seen from the smaller distance between the two peaks. XAS was used to confirm the 5C configuration of the substrate and cofactor bound sample, as the low energy transition could not be observed in the MCD spectrum.

6.2.2.1 Motivation for studies of TPHs electronic and geometric structure

PAH and TH have both been extensively studied, while results on TPH are very limited due to the inherent instability of the enzyme and thus the difficulty to purify it. The iron coordination of both PAH and TH in the active site has been studied using CD, MCD, and XAS(*67–70*), as described in the previous section. The studies showed that the resting state iron is 6 coordinated (6C) with 2 His, 1 Glu and 3 waters, but changes to 5C (2 His, bidentate Glu and 1 water) upon binding of both substrate and cofactor/cofactor analogue. This opens a coordination position for O_2 to react with iron. So far the active site of TPH has not yet been studied using CD and MCD, and even though the high sequence identity between the AAAHS suggests a similar mechanism, a recent crystal structure of TPH questions this. The crystal structure of doubly truncated *gallus gallus* TPH1 had the natural amino acid substrate bound and showed a 5C iron in the active site(*12*). However, an imidazole from the crystallization solution was also bound to iron, and it can be speculated whether this could mimic the function of the cofactor to induce the changes in iron coordination. Furthermore, the crystal structure contained Fe(III) and is thus not electronically consistent with the catalytically relevant species. In order to determine whether the binding of tryptophan alone triggers the iron coordination change from 6C to 5C, the active site of TPH should be analyzed with CD and MCD. Such a study would also elucidate whether TPH follows the general reaction mechanism of the other pterin-dependent hydroxylases.

6.2.3 Experimental procedures

All experiments except expression and purification were performed at Stanford University.

6.2.3.1 Protein expression and purification

The catalytic domain of human TPH2 (c*h*TPH2) was expressed and purified as previously described (see section 4.2.1 and 4.2.4.1). Protein purified for initial experiments were added 5 mM EDTA in order to remove iron from the active site. The apo-protein samples were primarily used for iron titration experiments (data not shown). All results featured in this chapter are made on protein samples that were saturated with iron during purification by addition of 1.5 molar equivalent of $(NH_4)_2Fe(II)(SO_4)_2$ before gel-filtration. Several purifications (~12) were made and the products from all purifications were mixed to achieve a homogeneous sample pool. A deuterated buffer was made by dissolving 20 mM HEPES and 300 mM NaCl in D_2O and adjusting pD to 7.2 with NaOD. To exchange hydrogens in the HEPES molecules for deuterium, the buffer was vacuum evaporated and subsequently dissolved in D_2O 3 times. The purified protein was exchanged into the deuterated buffer by 6 times ultrafiltration from 6 ml down to 1 ml following addition of 5 ml deuterated buffer.

SDS-PAGE and mass spectrometry was performed as described in section 4.2.5.1 and section 4.2.5.4, respectively. c*h*TPH2 concentrations between 1-10 µM in 0.6 M NH_4OAc were used. The spectrum of c*h*TPH2+Fe was smoothed using the Spectrum smooth option in the software suite MassLynx 4.1.

6.2.3.2 Preparation of CD and MCD samples

Solutions and solids were made anaerobic by purging with argon on a Schlenk line. Anaerobic reagents were transferred into a nitrogen atmosphere maintained in a Vacuum Atmosphere Nexus-1 glovebox (<5ppm oxygen). A fresh solution of sodium dithionite was made every day in deuterated buffer. Solutions of tryptophan were made in deuterated buffer, frozen and thawed each day before use. The BH_4 solution was made in a deuterated buffer with strong buffer capacity (100 mM HEPES) and subsequently pH was adjusted to 7.2 using NaOH in D_2O, as no NaOD was available. Dissolving BH_4 in the normal deuterated buffer without subsequent pH adjustment resulted in a very low pH, which caused protein precipitation upon mixing. The BH_4 samples were frozen and thawed each day before use.

Enzyme samples were made anaerobic by purging with argon. Inside the glovebox, 5 equivalents of sodium dithionite were added to the enzyme sample and it was incubated for 10 min while kept on a cold gel. 3 equivalents of tryptophan and/or BH_4 were then added and the samples were incubated for another 10 min. Samples for CD were injected into a 1 cm path length Infrasil quartz cuvette adapted with a ChemGlass stopcock to prevent air oxidation during the experiment. CD samples were kept on ice. Samples for MCD were added a glassing agent, which were either deuterated sucrose or deuterated glycer-3-ol. Deuterated sucrose was added to saturation, while deuterated glycer-3-ol was added to a concentration of 50 (v/v)%. The enzyme samples were injected into MCD cells, which consisted of two Infrasil quartz disks that were separated by a neoprene spacer (0.2-0.3 cm path length) and held in place by two fitted copper plates. After removal from the glovebox the MCD samples were directly frozen in liquid nitrogen and stored at 77 K.

The protein concentration varied in the samples between 1.0 and 2.0 mM. The natural substrate and cofactor were used for the experiments, which is only possible in an oxygen free environment, as the reaction would otherwise proceed.

6.2.3.3 Circular dichroism experiments

Near-IR CD spectra (278 K, 600-2000 nm) were obtained by using a Jasco J200D spectropolarimeter with a liquid nitrogen cooled InSb photodiode detector. The sample temperature was maintained by a recirculating water bath attached to the cell holder. Data acquisition was achieved using routines written within the software package LabVIEW (National Instruments). Contributions to the CD intensity due to buffer and cell backgrounds were subtracted from the protein CD spectra.

6.2.3.4 Magnetic circular dichroism experiments

Low-temperature near-IR MCD data (5-15 K, 1-7 T, 600-2000 nm) were collected on a Jasco J200D spectropolarimeter equipped with a liquid nitrogen cooled InSb photodiode detector and fitted with an Oxford Instruments SM 4000-7T superconducting magnet along the beam path. Each sample was measured at 7 T, 3.5 T, 0 T, -3.5 T, and -7 T and at two different temperatures (5 K and 15 K). For some conditions the scan rate was varied. MCD spectra were corrected for zero-field baseline effects either by subtracting the corresponding 0 T scan or by averaging the positive and negative field data (by subtracting the negative field data from the positive field data and dividing by 2). Gaussian fitting of the MCD spectra was performed using the program PeakFit v4.12 (SPSS Science).

6.2.4 Results and discussion

In this section the results from the CD and MCD measurement on c*h*TPH2 will be presented and discussed. It should be clarified here that this is the first time the active site of TPH is investigated using CD and MCD,

and the first time any CD and MCD study of the AAAHs uses the natural cofactor BH_4. The results have been divided into CD and MCD experiments. First, the purification product of c*h*TPH2 for the measurements is characterized.

The product from the c*h*TPH2 purifications was investigated by SDS-PAGE and mass spectrometry after exchange into the deuterated buffer. The results can be seen in Figure 97 and Figure 98. From the SDS-PAGE the sample looks pure, and it is thus suitable for use in CD and MCD measurements.

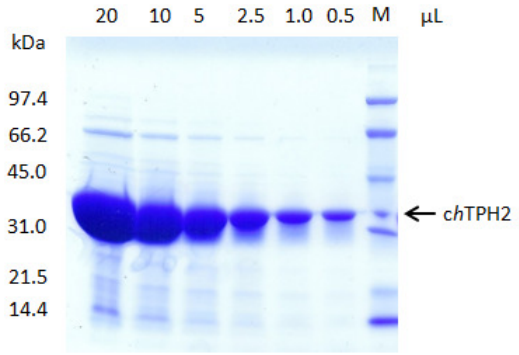


Figure 97: SDS-PAGE analysis of the purified c*h*TPH2 in deuterated buffer (diluted) for CD and MCD experiments. M: Molecular weight standard. Numbers: µL loaded

Mass spectra were taken of both c*h*TPH2 purified according to the general procedure, and c*h*TPH2 where either 5 mM EDTA or 1.5 molar equivalent $(NH_4)_2Fe(II)(SO_4)_2$ were added in the purification. When purification was performed according to the general procedure, c*h*TPH2 contains both the apo- and holo-form in a 40:60 ratio. The calculated masses of the apo and holo forms are 36325 Da and 36379 Da, respectively. The theoretical masses of apo and holo c*h*TPH2 are 36316.5 Da and 36369.3 Da (Fe(III)). The masses correspond quite well, however there is a constant difference of approx. 10 Da between the measured and theoretical masses. This could be caused by errors in the calibration. However, the mass difference between the apo and holo forms is 54 Da, which is in agreement with the theoretical mass of Fe(III) of 52.8 Da. Addition of 5 mM EDTA during purification effectively removes iron from the active site of c*h*TPH2, so the sample mainly contains the apo form. Addition of iron completely restores the holo form of c*h*TPH2. This last sample with added iron was used for all the CD and MCD measurements described in this chapter. It was decided to use the c*h*TPH2 sample that was saturated with iron during purification, as this eliminates the chance of any excess iron in the solution, which will disturb the active site iron signal in MCD measurements. As iron is added prior to gel-filtration, this purification step will remove the excess iron that has not been bound in the active site of c*h*TPH2. The earlier procedure for TH samples was to remove all iron from the enzyme sample after purification and subsequently titrate iron into the sample until approx. 85 % saturation was reached(*70*). However, this last option is experimentally quite difficult and requires very precise protein concentration measurements. Merely using a protein sample that already from purification is saturated with iron seems a much simpler way of avoiding excess iron in the sample.

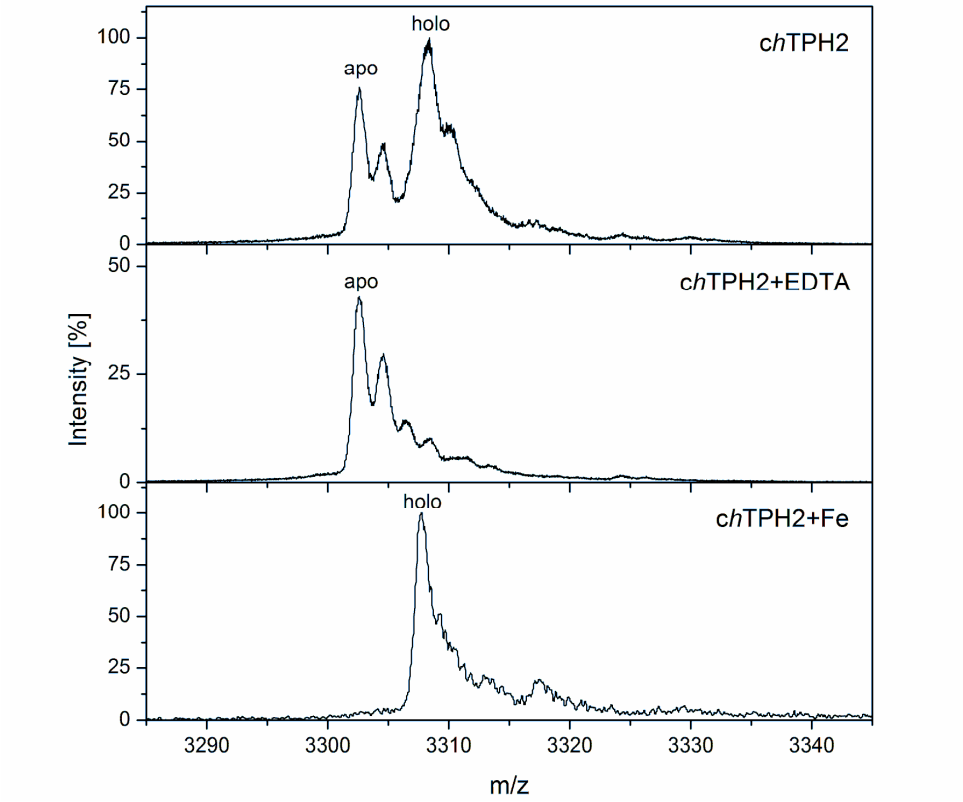


Figure 98: Mass spectra of *ch*TPH2 for MCD measurements. Top: *ch*TPH2 purified according to the general procedure. Center: *ch*TPH2 with 5 mM EDTA added in the purification. Bottom: *ch*TPH2 with 1.5 molar equivalent of $(NH_4)_2Fe(II)(SO_4)_2$ added in the purification.

6.2.4.1 Circular dichroism experiments

In this section the results from the CD experiments on the 4 substrate/cofactor combinations will be presented and discussed. The differences in intensities between samples should not be given any great importance, as the intensity depend on protein concentration, which is less accurately determined.

6.2.4.1.1 Glycerol samples

The 4 combinations (TPH-resting, TPH-trp, TPH-BH_4, and TPH-trp-BH_4) were measured in 50 (v/v)% deuterated glycer-3-ol at 5 °C and the results can be seen in Figure 99. The peak positions resulting from the Gaussian fitting can be seen in Table 23.

The spectrum of TPH-resting shows a positive peak around 8200 cm^{-1} and a negative peak around 11100 cm^{-1} indicative of a 6C coordination, where two peaks centered around ~10000cm^{-1} split by ~2000 cm^{-1} are expected(*160*). The split for the TPH sample is somewhat higher (~3000 cm^{-1}) indicating a loose 6th ligand, most likely water.

Addition of the cofactor BH_4 does not perturb the ligand field significantly since the peak positions are more or less identical to TPH-resting. This is also the case for PAH(*67*). When the substrate trp is added the peak pattern is the same, however the peaks shift to slightly lower energies. The split of the peaks is still ~3000 cm^{-1}. For PAH, binding of L-Phe also has an effect on the peak positions compared to the resting sample, however they move to higher energies(*69*).

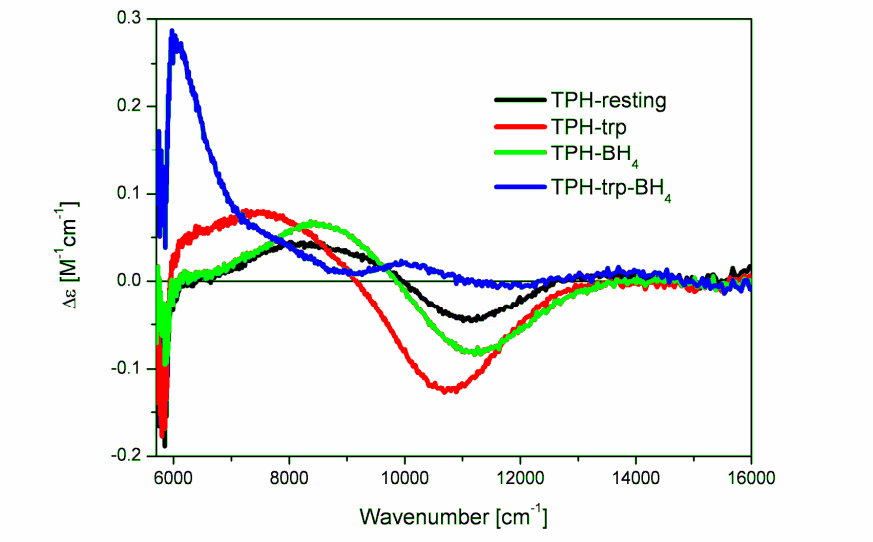


Figure 99: CD spectra for the 4 TPH substrate/cofactor combinations. The samples were all prepared with 50 (v/v)% deuterated glycer-3-ol as glassing agent.

When both trp and BH_4 are added, the peak pattern completely changes. A low energy peak appears, and 2 smaller peaks are seen at higher energies. The low energy peak is directly indicative of a 5C iron, which supports a 6C to 5C change of the active site iron only upon binding of both substrate and cofactor. A 5C square pyramidal site would give two transitions around 5000 and 10000 cm^{-1}, while a 5C trigonal pyramidal site theoretically gives transitions <5000 and <10000 cm^{-1}(*160*). As the third positive peak for TPH-trp-BH_4 is located around 10000 cm^{-1} this suggests a square pyramidal site, which is also consistent with the position of the low energy peak, that is clearly visible in the spectrum. For a 4C sample, two low energy peaks would be expected around 5-7000 cm^{-1}. The peak at 10000 cm^{-1} rules out this coordination. The presence of 3 peaks indicates a mixed sample, where not all TPH has been converted to the 5C form. Extra substrate and cofactor were added to the sample to test this, as described in the next section.

Table 23: Peak positions from the fitting of the CD spectra shown in Figure 99 for the 4 TPH substrate/cofactor combinations. The samples were all prepared with 50 (v/v)% deuterated glycer-3-ol as glassing agent.

Sample	Peak 1	Peak 2	Peak 3
TPH-resting	-	8245 cm^{-1}	11109 cm^{-1}
TPH-trp	-	7498 cm^{-1}	10758 cm^{-1}
TPH-BH_4	-	8534 cm^{-1}	11193 cm^{-1}
TPH-trp-BH_4	< 6000 cm^{-1}	7254 cm^{-1}	10008 cm^{-1}

6.2.4.1.2 Buffer samples

In order to check for any effect on the active site configuration of the glassing agent (deuterated glycer-3-ol) the 4 combinations were also made in deuterated buffer. The result is seen in Figure 100. Even though there are small changes in peak positions and intensities, the general picture and the trend are the same. It is especially important to notice that the samples with deuterated buffer also show a low energy peak when both trp and BH_4 are added. The spectra for the samples with deuterated buffer are smoother due to a stronger signal from the higher protein concentration of the undiluted sample.

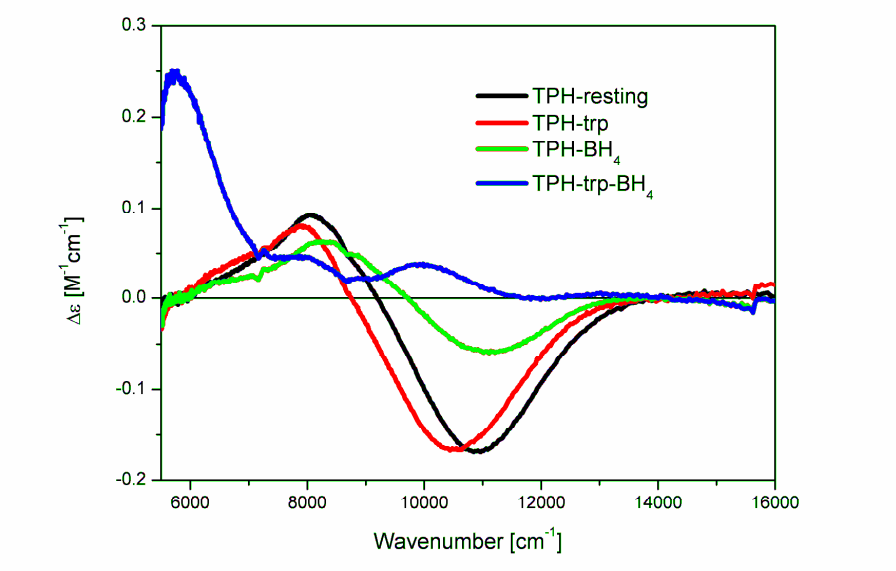


Figure 100: CD spectra for the 4 TPH substrate/cofactor combinations. The samples were all prepared in deuterated buffer with no glassing agent.

As the TPH-trp-BH_4 sample looks like a mixed sample, more substrate and cofactor were added to see if this had any effect on the spectrum. If the binding of trp or BH_4 is weak, a higher ratio than 1:3 might be needed to completely saturate the protein. The spectrum did not change upon addition of up to 9 equivalents of trp and BH_4 (data not shown).

6.2.4.1.3 Sucrose samples

From the MCD data (*vide infra*, see section 6.2.4.2) it was seen that there were differences between samples made in glycerol and sucrose, especially for combinations TPH-resting and TPH-trp. Therefore, these two combinations were made with both glycerol, sucrose and buffer to check for differences in the CD spectra. It is not possible to compare MCD data of the buffer sample, since a glassing agent is needed to prevent icing. Thus, CD data often is the best way to determine which of the glassing agents that best mimic the natural state of the protein in buffer.

Generally, the spectra for all 3 samples have the same features; A positive peak around 8000 cm^{-1} and a negative peak around 11000 cm^{-1} (see Figure 101). However, there are small differences in the shape of the curves, especially for TPH-resting sample. The sucrose sample has a kink around 9000 cm^{-1} not seen for the other two samples. Based on the spectra shown here it is not possible to determine which of the two glassing agents best represent the natural state of the protein. The effect of glycerol and sucrose on the active site will be discussed further in section 6.2.4.2 and 6.3.

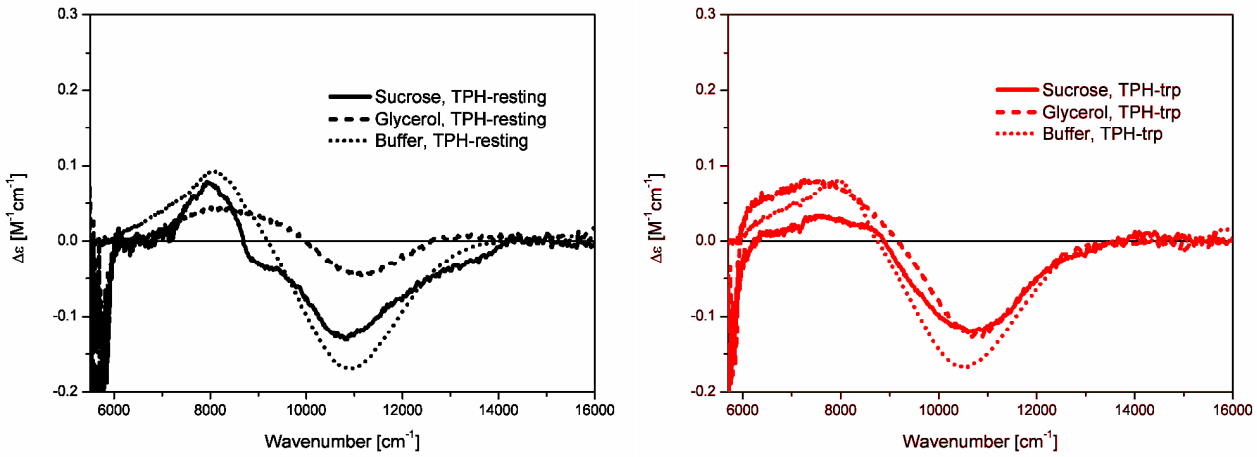


Figure 101: CD spectra of the TPH-resting sample (left in black) and TPH-trp sample (right in red). The full line is the sample with deuterated sucrose as glassing agent, the dashed line is the sample with deuterated glycer-3-ol as glassing agent, and the dotted line is the sample in deuterated buffer with no glassing agent.

6.2.4.2 Magnetic circular dichroism experiments

In this section the results for the magnetic circular dichroism experiments will be presented. If nothing else is mentioned, the spectra are all recorded at 5 K with a magnetic field of 7 T.

6.2.4.2.1 Glycerol samples

The MCD spectra were first recorded with 50 (v/v)% deuterated glycer-3-ol as the glassing agent for the 4 TPH substrate/cofactor combinations (TPH-resting, TPH+trp, TPH+BH_4 and TPH+trp+BH_4). The result is seen in Figure 102. Compared to the MCD data for PAH and TH in the literature (see Figure 96), the TPH data are of very high quality including strong signals with minimal noise. This can be attributed to both pure protein samples of high concentrations and effective exchange of hydrogen for deuterium.

TPH-resting shows two clearly separated peaks at around 8500 cm^{-1} and 10800 cm^{-1}, which corresponds well with a 6 coordinated (6C) iron and also is similar to corresponding samples for TH (8720 and 10750 cm^{-1})(*70*) and PAH (8500 and 10300 cm^{-1})(*69*). When tryptophan binds the ligand field is slightly perturbed but iron remains 6C. This is seen from the two peaks at 8200 and 10000 cm^{-1}. The MCD data cannot be fitted with one peak and the CD data also supports the 6C as the two peaks can be clearly distinguished (one positive and one negative). Thus, binding of tryptophan alone does not trigger a change in iron coordination to 5C, which has been speculated based on the crystal structure of *gallus gallus* TPH with tryptophan bound. For TPH-BH_4 the two peaks from the CD spectra are not obvious in the MCD spectra of the same sample. One peak at 9200 cm^{-1} is clearly seen, however the high energy peak apparently becomes negative, and is hard to distinguish. As the CD spectrum clearly shows two peaks, it is concluded that binding of BH_4 to TPH does not change the coordination of iron significantly as it remains 6C. The major change in the spectrum is seen when both tryptophan and BH_4 are added. A low energy peak emerges at <6000 cm^{-1} corresponding with the change in iron coordination to 5. At higher energy two peaks are seen at 8700 and ~10000 cm^{-1}. The two peaks at higher energy could be explained if the sample contains a mixture of iron 5C and 6C. This is the first time that MCD data of any AAAH with the natural cofactor BH_4 has been presented. Earlier studies have used cofactor analogues such as the redox inactive 5-deaza-6-MPH_4 or 6MPH_4.

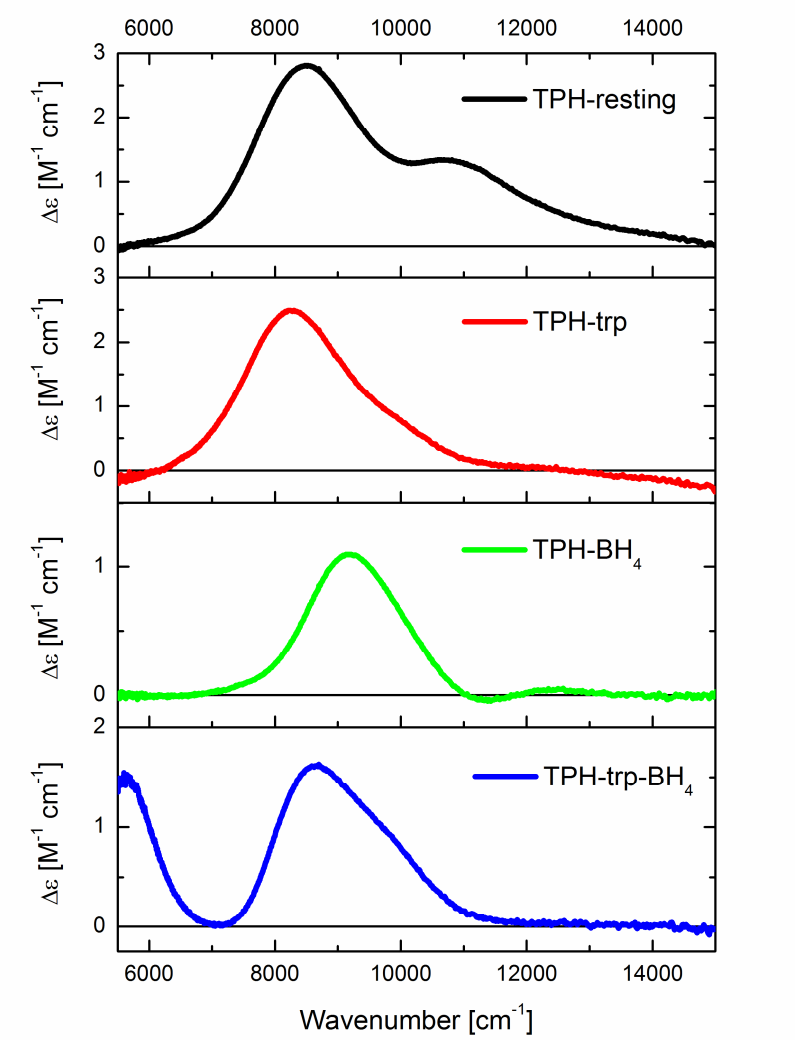


Figure 102: MCD spectra for the 4 TPH substrate/cofactor combinations. The samples were all prepared with 50 (v/v)% deuterated glycer-3-ol as glassing agent.

The small negative peak of the TPH-BH_4 sample around 11200 cm^{-1} was subject to further investigations, as it is not clear whether it represents an actual peak (or maybe one negative and one positive peak). This was investigated by recording spectra at different strengths of magnetic field and different temperatures. The result in seen in Figure 103, and does not clearly show whether the peaks are $\mathscr{C}$-term or just baseline problems. As described earlier $\mathscr{C}$-term transitions are both field and temperature dependent. The relatively weak signal of the TPH-BH_4 sample made interpretation even harder. However, as the CD spectrum of the TPH-BH_4 sample shows a negative peak at ~11200 cm^{-1} corresponding more or less to the position of the small negative peak in the MCD spectra, it is likely that it represents an electronic transition.

It was attempted to fit the peak positions and widths from both CD and MCD spectra combined, but this did unfortunately not succeed. The MCD data generally shows a smaller splitting between the two excited states than the CD data for the TPH-trp and TPH-BH_4 samples.

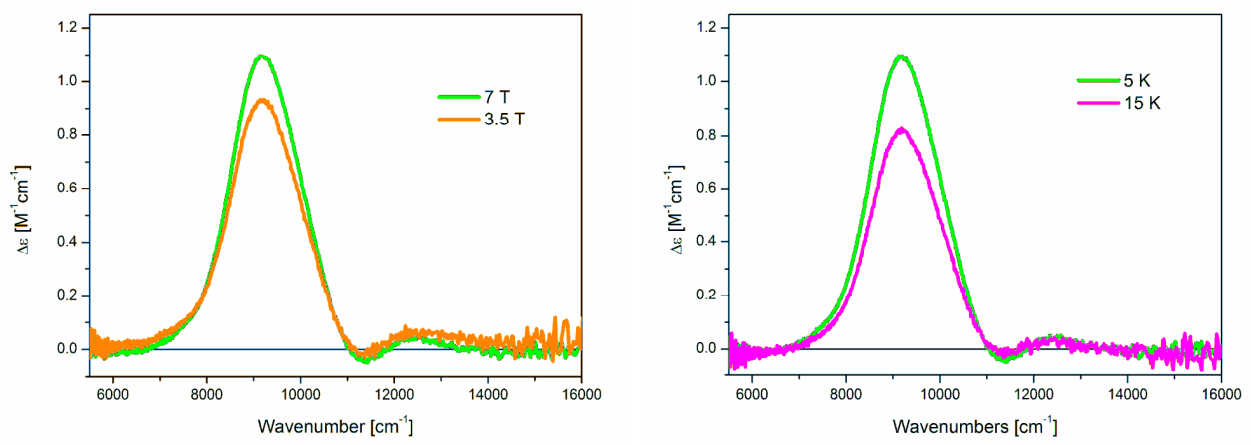


Figure 103: MCD spectra of the TPH-BH_4 sample. The sample was prepared with 50 (v/v)% deuterated glycer-3-ol as glassing agent. Left: The field effect. Right: The temperature effect.

6.2.4.2.2 Sucrose samples

In order to check whether glycerol has an effect on the active site geometry, another glassing agent was tested. Deuterated sucrose was chosen as this was standard in the Solomon Laboratory, and as sucrose usually gives a good glass in the sample upon freezing. The results of this test with sucrose were surprising, and can be seen in Figure 104. The overall trend in the experiment was the same as for glycerol samples; iron changed from 6C to 5C only upon binding of both tryptophan and BH_4, shown by emerge of a low energy peak. However, there were clear changes in the ligand fields around iron as compared to the glycerol samples for all samples except TPH+trp+BH_4. For TPH-resting in sucrose the peaks shift further away from each other, and the highest energy peak is almost indistinguishable. This could lead to some doubt on whether the TPH-resting sucrose sample is 6C, but variable temperature variable field MCD measurements performed by graduate student Lei V. Liu confirm that the coordination is 6C (data not shown). Furthermore, the CD spectrum of the TPH-resting sample in sucrose also displays two peaks with positions indicative of a 6C coordination. For TPH+trp the high energy peak becomes negative instead of positive and for TPH+BH_4 the intensity relationship between the two peaks shifts.

These data show that there is a difference in the iron coordination sphere depending on the glassing agent. It is hard to distinguish whether glycerol or sucrose binds directly to the iron or changes the ligand field by affecting/pushing the iron coordinated waters. Apparently, the glassing agents can only affect the active site in the open 6C conformation, as the TPH-trp-BH_4 sample is not sensitive to the type of glassing agent. It can be speculated that when the active site of TPH is 6C and open, the glassing agents are free to coordinate to iron in competition with water molecules. However, when the coordination changes to 5C, the active site closes and does not allow any interaction between the glassing agents and iron, which is no longer freely accessible.

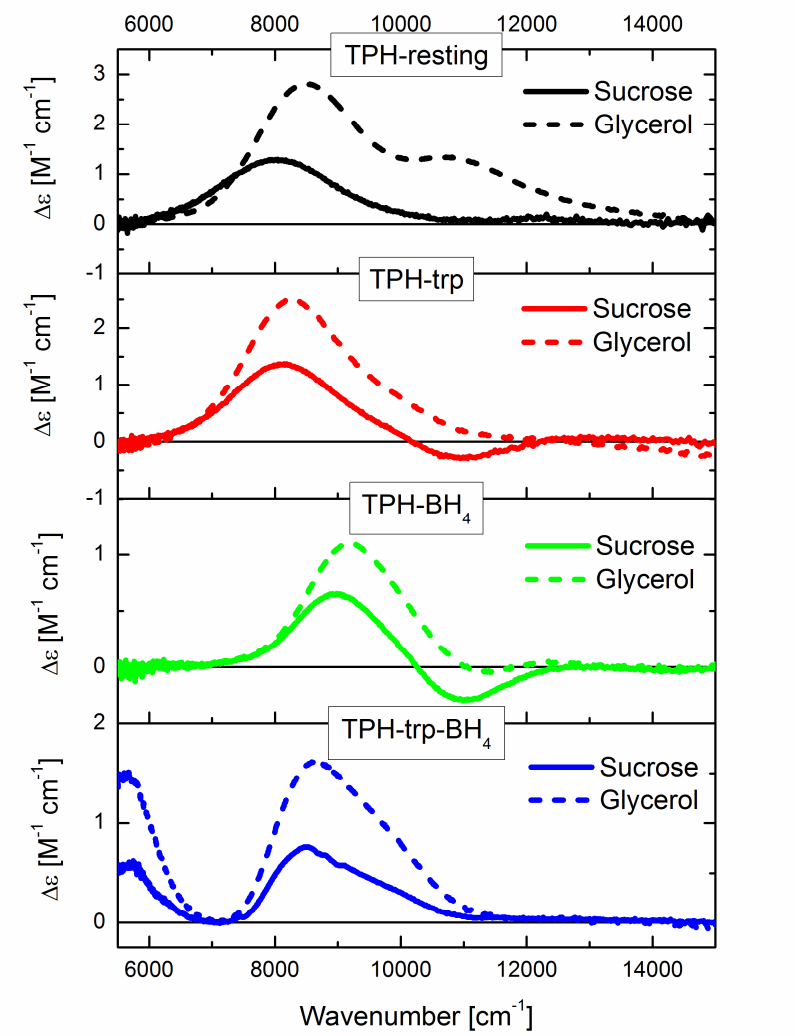


Figure 104: MCD spectra for the 4 TPH substrate/cofactor combinations. Full lines are samples prepared with deuterated sucrose as glassing agent. Dashed lines are samples prepared with 50 (v/v)% deuterated glycer-3-ol as glassing agent.

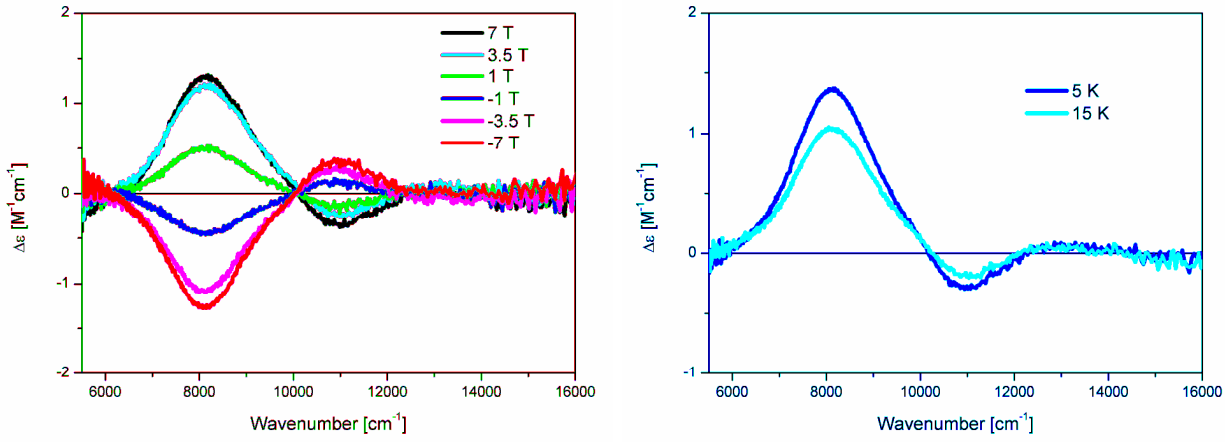


Figure 105: MCD spectra of the TPH-trp sample. The sample was prepared with deuterated sucrose as glassing agent. Left: The field effect. Right: The temperature effect.

The TPH-trp sucrose sample was investigated further for field and temperature effect, to clearly distinguish the high energy negative peak. From Figure 105 it is seen that the peak does indeed represent the $\mathscr{C}$-term as it shows both field and temperature dependence.
For the TPH-BH_4 sucrose sample, both the negative and positive peak is also field dependent as seen in Figure 106.

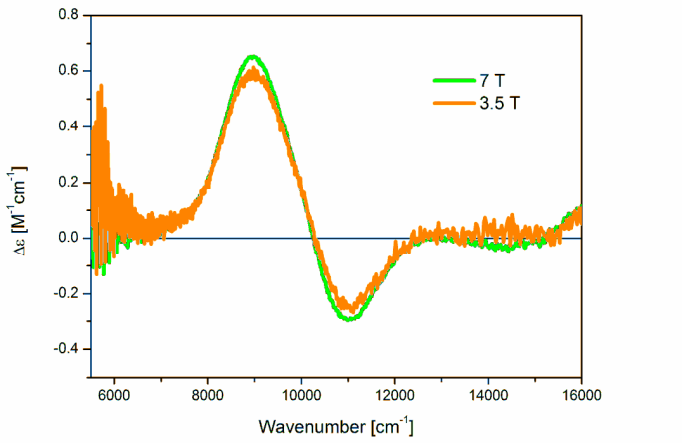


Figure 106: MCD spectra of the TPH-BH_4 sample. The sample was prepared with deuterated sucrose as glassing agent. The spectra demonstrates the field effect.

The peak positions of the 4 TPH substrate/cofactor combinations in both glycerol and sucrose can be seen in Table 24. There are significant differences between similar samples in different glassing agents, except for the TPH-trp-BH_4 sample.

Table 24: Results from the Gaussian fitting of the MCD data for samples with 50 (v/v)% deuterated glycer-3-ol or deuterated sucrose as glassing agent. TPH-trp-BH_4 was fitted with both 3 and 4 peaks, and both results are shown in the table. *: Negative peak. (): insignificant peak.

Sample	Glassing agent	Peak 1	Peak 2	Peak 3	Peak 4
TPH-resting	Glycerol	-	8383	10397	-
	Sucrose	-	8038	(11560)	-
TPH-trp	Glycerol	-	8182	9603	-
	Sucrose	-	8164	*11006	-
TPH-BH_4	Glycerol	-	9243	*(11228)	-
	Sucrose	-	9004	*11007	-
TPH-trp-BH_4	Glycerol (3 peaks)	~5531	8447	9412	-
	Glycerol (4 peaks)	~5903	8264	8804	9591
	Sucrose (3 peaks)	~5814	8370	9321	-
	Sucrose (4 peaks)	~5628	8285	8896	9648

Comparison of CD results from samples with no glassing agent added and those with glycerol and sucrose gave no convincing results for which structures best represent the natural state (see section 6.2.4.1.3).

6.2.5 Conclusions

The 4 TPH substrate/cofactor combinations were all investigated by CD and MCD. From CD experiments, a clear low-energy peak corresponding with a change in iron coordination from 6 to 5 is seen only when both substrate and cofactor are added. The same is the case for the MCD experiments. It can thus be concluded that the coordination of iron in TPH is 6C in the resting state and only changes to 5C when both substrate and cofactor bind. Small changes in the ligand field were seen when trp or BH_4 binds separately, but iron remains 6C. The results obtained here are similar to results from PAH and TH supporting the assumption that they all share a similar reaction mechanism, in which the conformational change from open to closed upon substrate and cofactor binding is accompanied by a change in iron coordination from 6C to 5C. The binding of tryptophan does not lead to a change in iron coordination to 5C, as was otherwise implied by the recent crystal structure of *gallus gallus* TPH1 with tryptophan bound.

This is the first time a MCD study has been performed on the AAAHs using the natural substrate and the natural cofactor. The data quality is very high and TPH is thus a promising candidate for further investigations of the AAAHs reaction mechanism.

Deuterated sucrose and glycerol were tested as glassing agents, and even though TPH samples with both glassing agents showed the same overall trend, there were differences in the spectra indicating that either glycerol, sucrose or both has an effect on the active site iron. This is investigated further by other techniques in the next section.

6.3 The glycerol effect

Indications from MCD experiments lead to further investigations of the effect of glycerol on the active site of c*h*TPH2. The results from these will be presented in this section.

6.3.1 Experimental procedures

6.3.1.1 Purification

chTPH2 was purified according to the described procedures (see section 4.2.4.1) either in buffer with 5 (w/v)% glycerol or without glycerol.

6.3.1.2 Absorbance measurements

The absorbance of c*h*TPH2 samples with and without glycerol, and with glycerol added in concentrations ranging from $5 \cdot 10^{-7}$ (v/v)% to 20 (v/v)% was measured using a NanoDrop spectrophotometer. The corresponding buffer was used as blank and at least 2 consecutive measurements were performed for each sample. 0.05 M sodium dithionite was also added to the purification products with and without glycerol, and the absorbance before and after was measured.

6.3.1.3 Iron and glycerol interactions

To test the effect of glycerol on Fe(II) and Fe(III) in aqueous solution, the following experiment was conducted. 10 mM solutions of $(NH_4)_2Fe(II)(SO_4)_2$ and $NH_4Fe(III)(SO_4)_2$, respectively were made in MilliQ water. 20 (v/v)% glycerol was added to the solutions and the color change was observed.

6.3.1.4 Activity assay

The activity of c*h*TPH2 with different glycerol concentrations were tested using the general activity assay described in section 4.2.5.2. The protein concentration was 0.5 µM and the temperature was 27-28 °C due to the lack of cooling equipment. Each sample was measured 8 times.

6.3.2 Results and discussion

In the following section, the results from the experiments on the effect of glycerol on the active site of TPH will be presented and discussed.

6.3.2.1 Purification

Overall, the chromatograms for the purifications with and without glycerol were similar, except in the gel-filtration step, where a peak at 416 nm is seen when c*h*TPH2 elutes when using buffer with 5 % glycerol, and no peak at 416 nm is seen when using buffers without glycerol. This gave an indication that the difference between c*h*TPH2 samples with and without glycerol can be observed by absorbance at 416 nm corresponding to a yellow color. Normally, purifications of TPH are monitored at 280 nm (aromatic amino acids), 416 nm (possible heme-containing contaminants) and 214 nm (peptide bonds), which is why this wavelength was chosen for the purification. Upon concentration of the two purification products, a visible color difference was also seen. The sample with 5 (w/v)% glycerol was clearly more yellow than the sample containing no glycerol, see Figure 107. The yields of the purifications with and without glycerol were similar.

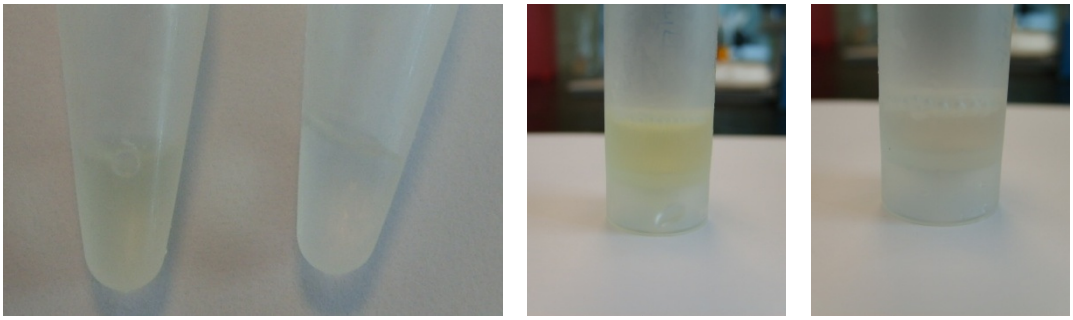


Figure 107: Color difference between c*h*TPH2 sample with and without 5(w/v) % glycerol. Left: Eppendorf tubes with glycerol (left, yellow) and no glycerol (right, colorless) c*h*TPH2 solution with the same concentration. Center: Concentrated 5 (w/v)% glycerol c*h*TPH2 solution with a clear yellow color. Right: Concentrated c*h*TPH2 solution with no glycerol and an orange/brown color.

6.3.2.2 Absorbance measurements

The purifications with and without glycerol showed a difference in absorbance at 416 nm. Thus, the following analysis also used this wavelength to investigate the effect of glycerol on the protein samples. The absorption spectra for two samples with the same protein concentration, but with 0 % and 10 % glycerol, respectively, are shown in Figure 108. There is a clear difference in the absorption between the two samples from approx. 310 nm to 500 nm. Absorption in the range 420-430 nm would give a yellow substance, corresponding with the yellow color of the glycerol containing TPH sample.

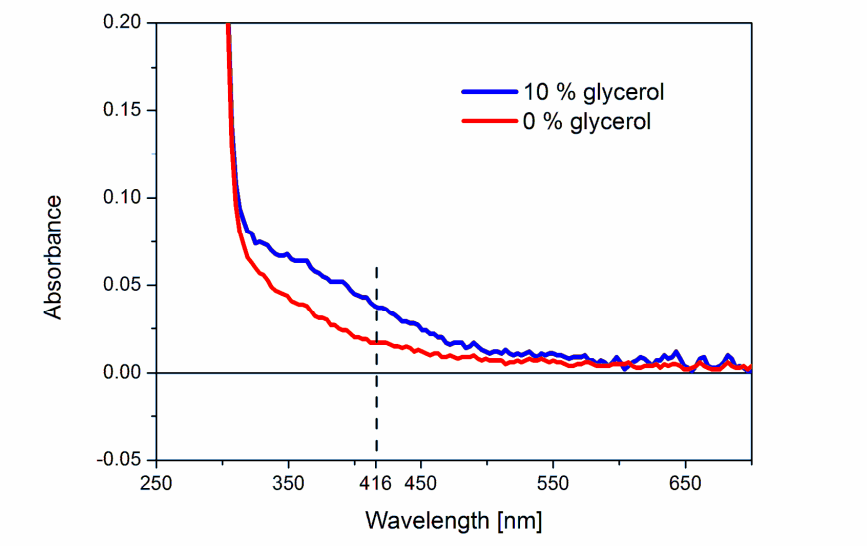


Figure 108: Absorbance spectra of two *ch*TPH2 samples without (red) and with 10 (v/v)% glycerol (blue). The dashed line marks 416 nm. The two samples have the same protein concentration.

The two concentrated products from the purifications were added glycerol in different concentrations and the absorption at 416 nm was measured. The results can be seen in Table 25. For the sample with 5 (w/v)% glycerol, no increase in absorption upon addition of extra glycerol is seen. The start absorption is however relatively high compared to the sample without glycerol. For the sample with no glycerol in the buffer, the results show a clear increase in absorbance at 416 nm upon addition of glycerol. Higher concentration of glycerol does not increase the absorption further. The absorptions measured were however very low, so the sample without glycerol was concentrated and measured again. The increase in absorption at 416 nm upon addition of glycerol is also seen in the concentrated sample.

Table 25: Absorbance at 416 nm for samples of *ch*TPH2 upon addition of different glycerol concentrations. Each sample was measured at least 3 times and the standard deviation is given.

(v/v)% glycerol added	Abs. 416 nm ± St. Dev. 230 µM *ch*TPH2 5 (w/v)% glycerol	Abs. 416 nm ± St. Dev. 182 µM *ch*TPH2 No glycerol	Abs. 416 nm ± St. Dev. 590 µM *ch*TPH2 No glycerol
0	0.024 ± 0	0.006 ± 0.001	0.016 ± 0.002
5	0.027 ± 0.001	0.018 ± 0.001	
10	0.028 ± 0.002	0.018 ± 0.001	0.034 ± 0.005
20	0.027 ± 0.002	0.017 ± 0	

To investigate how low a concentration of glycerol that would still cause the increase in absorption, two new experiments were set up with very low glycerol concentrations. The results can be seen in Table 26, and show somewhat ambiguous results. The 1st experiment shows an increase in absorption at the lowest glycerol concentration of $5 \cdot 10^{-5}$ (v/v)%, which corresponds to approx. 7 µM. In the 2nd experiment, the absorption does not increase until a glycerol concentration of 0.5 (v/v)% corresponding to approx. 70000 µM. The goal of these experiments was to determine if glycerol was binding 1:1 with the protein suggesting a specific binding, however the data are not clear on this.

Table 26: Absorbance at 416 nm for samples of *ch*TPH2 upon addition of low concentrations of glycerol. Each sample was measured at least 2 times and the standard deviation is given, except for the sample marked with *, which was only measured one time.

(v/v)% glycerol	1st experiment Abs. 416 nm ± St. Dev. 437 µM *ch*TPH2	2nd experiment Abs. 416 nm ± St. Dev. 411 µM *ch*TPH2
0	0.028 ± 0.002	0.018 ± 0.003
$5 \cdot 10^{-7}$		0.019 *
$5 \cdot 10^{-5}$	0.051 ± 0.01	0.017 ± 0
$5 \cdot 10^{-4}$	0.051 ± 0.002	
$5 \cdot 10^{-3}$	0.047 ± 0.003	0.019 ± 0.001
$5 \cdot 10^{-2}$	0.049 ± 0.004	
0.5	0.060 ± 0.006	0.029 ± 0.001
1	0.053 ± 0.003	
2	0.057 ± 0.006	
5	0.051 ± 0.002	0.033 ± 0.001
7.5	0.055 ± 0.001	
10	0.050 ± 0.004	

The two concentrated products from the purifications were added dithionite to see if any effect on the absorbance at 416 nm was seen. As seen from Table 27, there was a clear decrease in the absorbance at 416 nm for the sample with 5 (w/v)% glycerol, whereas the dithionite had no effect on the sample without glycerol.

Table 27: The absorbance at 416 nm for samples of *ch*TPH2 upon addition of 0.05 M sodium dithionite. Each sample was measured 2 times.

	Abs. 416 nm ± St. Dev. 408 µm *ch*TPH2 5 (w/v)% glycerol	Abs. 416 nm ± St. Dev. 315 µM *ch*TPH2 No glycerol
Before dithionite addition	0.049 ± 0.001	0.013 ± 0.001
After dithionite addition	0.004 ± 0.001	0.018 ± 0.001

The addition of dithionite to the glycerol containing sample was also followed by a clear color change from yellow to colorless. These results suggest that only Fe(III) shows the yellow color in combination with glycerol, while the reduced Fe(II) is colorless even in the presence of glycerol. From these experiments it cannot be determined whether the disappearance of the yellow color is due to the reduction of iron or because glycerol does not bind to Fe(II).
Experiments with addition of Fe(III) and EDTA gave no conclusive results (data not shown).

6.3.2.3 Iron and glycerol interactions

To test the effect of glycerol on Fe(II) and Fe(III) in aqueous solution, solutions of $(NH_4)_2Fe(II)(SO_4)_2$ and $NH_4Fe(III)(SO_4)_2$ were made in MilliQ water. The Fe(II) solution was colorless, while the Fe(III) solution was red/orange. Upon addition of 20 (v/v)% glycerol the Fe(II) solution remained colorless, while the Fe(III) solution changed color from red/orange to yellow/orange. The colors observed in this simple test

corresponded to the colors of the concentrated protein solutions with glycerol (yellow) and without glycerol (brown/red), see Figure 107.

6.3.2.4 Activity assay

If glycerol binds in the active site it could have an effect on the reaction rate. Thus, the initial rate was measured for c*h*TPH2 samples with different concentrations of glycerol using the activity assay.

Table 28: Activity of c*h*TPH2 samples with different glycerol concentrations. The concentration of c*h*TPH2 in the assay was 0.5 µM. Each sample was measured 8 times. The unit for activity [int/min] is an arbitrary unit, but can be used for comparison between different samples.

(w/v)% glycerol	Activity ± St. Dev. [int/min]
0	515 ± 56
0.4	545 ± 36
0.8	567 ± 14

The results from the activity measurements in Table 28 show that glycerol does not inhibit the hydroxylation reaction of TPH. Actually, it slightly increases the activity and significantly decreases the standard deviations. This is in agreement with the fact that glycerol is a known protein stabilizer(*166*), and that a more stable protein solution would give lower standard deviations.

6.3.2.5 Theoretical calculations

In an attempt to specify which of the two glassing agents used in MCD that affects the active site, a theoretical calculation study was carried out by master student Niels Theis Bendtsen under supervision of Associate Professor Günther H. J. Peters(*167*). A model of the binding pocket of TPH2 was computed based on the 1MLW structure of c*h*TPH1 by point-mutating three amino acids. Several programs were used to dock sucrose and glycerol into the binding pocket and investigate their interaction with the active site iron. The calculations showed that both sucrose and glycerol can enter the active site of TPH and coordinate to iron.

6.3.3 Overall discussion

It is well known that catecholamines such as dopamine bind directly to iron in PAH(*129*) and TH(*168*) and inhibit catalysis. A crystal structure of PAH with bound dopamine shows a bidentate coordination to iron through two OH-groups(*129*). Glycerol could bind to iron in a similar binding mode. The light yellow color of the glycerol Fe(III)-TPH sample is also seen for Fe(III)-PAH in glycerol and disappears upon reduction(*69*). Furthermore, a glycerol effect was observed in some of the first PAH MCD experiments for the PAH sample with added tryptophan. The presence of tryptophan perturbed the iron active site, but upon addition of glycerol this effect was nullified and the peaks returned to the positions for the resting enzyme(*69*). It was speculated that glycerol displaces tryptophan from the active site pocket. However, the resting and L-phe bound PAH enzyme samples are insensitive to glycerol addition(*69*). Fe(III) in aqueous solution with glycerol is light yellow, suggesting that the color of the TPH samples arises from a direct interaction between iron and glycerol, and not from distant binding of glycerol in the active site. Glycerol is a smaller molecule than sucrose and it could more easily enter the active site, even though the theoretical calculations show that sterically both molecules fit.

6.3.4 Conclusions

There is a clear color change to yellow when glycerol is added to a solution of c*h*TPH2 in the Fe(III) form. The color disappears upon reduction to Fe(II) with sodium dithionite. Based on this it is reasonable to believe that glycerol is affecting the iron coordination, and that the sucrose samples used in MCD experiments most likely best represents the natural state of c*h*TPH2. Glycerol is commonly used in preparation of AAAH samples and it is thus an important finding that it interacts with the active site iron. However, as activity is not inhibited by glycerol, it may not have large consequences for the kinetic analysis. Glycerol is a known protein stabilizer and has been used in protein preparation for many years.

6.4 Differential scanning fluorimetry

The unfolding/melting temperature of a protein can be measured by DSF as described in section 4.3.2.3, and if binding of substrates thermodynamically stabilizes the protein, this should be visible from a DSF experiment. Different substrates and substrate analogous were added to c*h*TPH2 and ct*h*TPH2 and the melting temperature was determined.

6.4.1 Experimental procedures

chTPH2 and ct*h*TPH2 were purified as described in section 4.2.4.1 and 4.2.4.2. The samples were frozen in liquid nitrogen and stored at -80 °C. DSF experiments were performed at MaxLab, Lund University following the general procedure described in section 4.2.5.3. The protein concentration was 2 µM in the final screen, and the protein was diluted in 10 mM HEPES/NaOH, pH 7.0 before addition to the screen. For 1000 µL diluted protein solution, 3.0 µL 5000x SyPro Orange was used. The reference buffer was 100 mM HEPES/NaOH, 100 mM $(NH_4)_2SO_4$, pH 7.0.

The screen used was homemade, and was prepared by adding 1.5 mL reference buffer to each well in the deep well block. Substrates were added to the deep well block from frozen stocks solutions on the day of use, before the Tecan robot sat up the DSF screen. The final substrate concentrations were 100 µM. Iron-compounds were dissolved immediately before use to 66 mM stock solutions and diluted if necessary. 1 µL iron-compound solution per well was added directly to the DSF screen right before protein addition, in order to prevent oxidation. 8 µL protein/SyPro orange solution was added to each well with 24 µL screen (and 1 µL iron-compound).

6.4.2 Results and discussion

Initial measurements showed a large increase in melting temperature (5 °C) upon addition of $(NH_4)_2Fe(II)(SO_4)_2$ (data not shown). The effect of different $(NH_4)_2Fe(II)(SO_4)_2$:protein ratios were thus tested along with addition of Fe(III) as $NH_4Fe(III)(SO_4)_2$ and Fe(III)OAc. The results can be seen in Figure 109.

Addition of Fe(III) as $NH_4Fe(III)(SO_4)_2$ and Fe(III)OAc has no significant effect on the stability in terms of unfolding temperature. Mass spectrometric studies have shown that c*h*TPH2 is fully loaded with iron after purification (see section 6.2.4), so addition of extra Fe(III) does not make any difference. A significant increase in melting temperature is however observed with increasing $(NH_4)_2Fe(II)(SO_4)_2$ concentration. It must be assumed that TPH in solution contains Fe(III), as Fe(II) would have been oxidized by oxygen in the air and buffers. Thus, it makes sense that addition of Fe(II), which is the catalytically active species, stabilizes the protein relative to Fe(III). The constant increase in melting temperature with Fe(II):protein ratios up to 30 could be explained by a low binding affinity for Fe(II), when Fe(III) is already bound, or by the slow electron transfer between Fe(II) in solution and Fe(III) bound in the protein. The finding that TPH in

the Fe(II) form is significantly more stable than in the Fe(III) form is important and should be considered for e.g. crystallization and other characterization techniques.

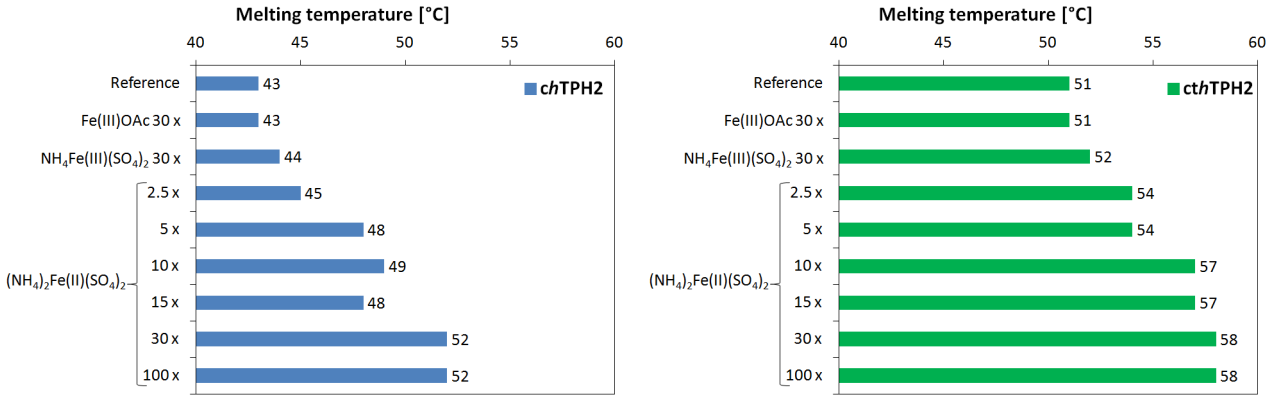


Figure 109: Results from DSF studies on the effect of addition of iron. Reference is 100 mM HEPES/NaOH, 100 mM $(NH_4)_2SO_4$, pH 7.0. Numbers (e.g. 30 x) refers to excess of the compound relative to the protein concentration, which was 2 µM. Left in blue: *ch*TPH2. Right in green: ct*h*TPH2.

Several types of reducing conditions were tested to investigate if the stabilizing effect could be achieved by reducing protein bound Fe(III) to Fe(II). No conclusive results were however seen, probably due to the negative effect of reducing agents on protein stability.

Addition of substrates (trp, BH_4), substrate analogue (BH_2) and possible regulators (serotonin) to TPH gave no significant increase in melting temperature (max 1 °C, data not shown). Addition of BH_4 gave a significant decrease in melting temperature probably caused by a low pH in the added BH_4 solution. Thus, BH_4 is not included in the following figures.

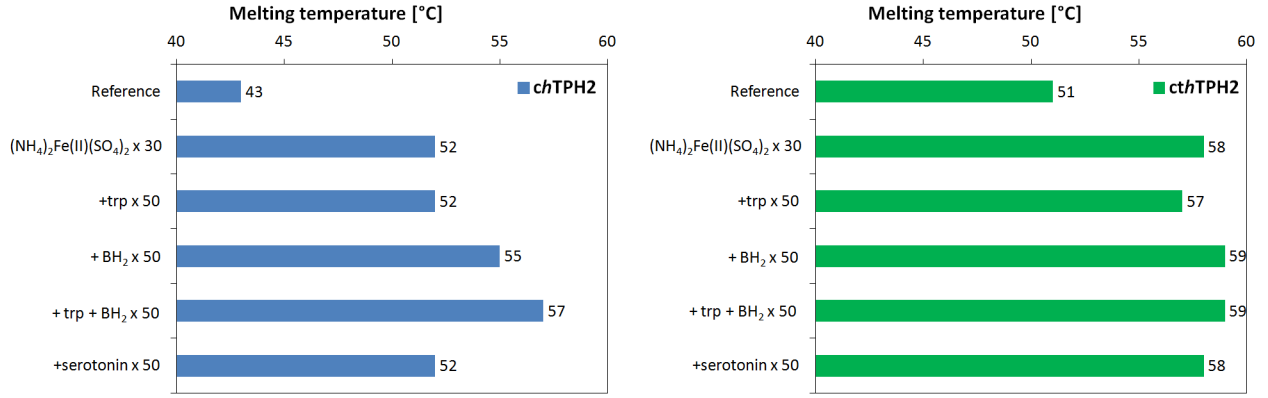


Figure 110: Results from DSF studies on the effect of addition of substrates to samples with Fe(II) added. Reference is 100 mM HEPES/NaOH, 100 mM $(NH_4)_2SO_4$, pH 7.0. Numbers (e.g. x 30) refers to excess of the compound relative to the protein concentration, which was 2 µM. All conditions with added substrates also contained $(NH_4)_2Fe(II)(SO_4)_2$ x 30. Left in blue: *ch*TPH2. Right in green: ct*h*TPH2.

However, if substrates are added along with Fe(II) an effect is seen on the melting temperature for *ch*TPH2, see Figure 110. Addition of BH_2, which is a cofactor analogue and inhibitor, gives a 3 °C increase in melting temperature, while addition of both trp and BH_2 gives an increase of 5 °C. Addition of trp alone does not have any effect. This suggests that BH_2 and trp binds to the protein and energetically stabilizes it. Furthermore, it indicates that even if tryptophan can bind before BH_2, it is energetically more favored when

binding after BH_2. This stabilizing effect could be caused by the large structural changes and the closing of the active site upon cofactor and substrate binding. Furthermore, these results suggests that binding of substrates in the Fe(III) form does not stabilize the protein. For ct*h*TPH2 the effect of substrates on the melting temperature is not significant.

6.4.3 Conclusions

Addition of Fe(II) increases the unfolding temperature of both c*h*TPH2 and ct*h*TPH2 with 9 and 7 °C, respectively. Addition of BH_2 and trp+BH_2 significantly stabilizes c*h*TPH2 in the presence of Fe(II). Without addition of Fe(II) no effect of BH_2 and trp+BH_2 addition is seen. This suggests that binding of the substrates only stabilizes the structure in the catalytically active state.

6.5 Isothermal titration calorimetry

The binding of substrates can be studied using a steady state kinetic assay and michaelis menten kinetic theory. For TPH, several studies using kinetics comparing K_M values and the differences between the isoforms have been published(*13*, *34*, *51*, *98*, *102*, *169*, *170*). In 2009, the Metalloprotein Chemistry and Engineering group investigated the kinetic differences between purified c*h*TPH1 and c*h*TPH2 at the exact same conditions, so the results can be directly compared(*13*). Despite the large sequence identity between the two catalytic domains of 81 %, several differences were seen. First, c*h*TPH1 was inhibited at large concentrations of tryptophan, while c*h*TPH2 was not. Second, the K_M value for BH_4 was more than 10 times larger for c*h*TPH1 than for c*h*TPH2. The latter is surprising, as none of the amino acids close to the BH_4 binding pocket differs between the isoforms. The results are in contrast to earlier studies of full-length TPH1 and TPH2(*51*), where the K_M values for BH_4 and Trp were in the same range for both isoforms. However, this study was performed on affinity-tagged TPH variants that were not completely pure.

As K_M values are difficult to compare and represent several factors beside the actual binding affinity, another technique was used to study the binding of substrate and cofactor to TPH. The technique is called isothermal titration calorimetry, and allows for direct determination of dissociations constants, which can be directly compared.

Isothermal titration calorimetry (ITC) is a technique used to determine the thermodynamic parameters of biomolecular interactions in solution. For ligands binding to protein, both the binding affinity, number of binding sites, entropy and enthalpy of the reaction are determined in a single experiment(*171*).

A representation of a typical ITC instrument can be seen in Figure 111. The ligand solution is placed in a syringe and added in small portions to the sample cell containing protein solution. The reference and sample cells are mounted in an adiabatic environment and can only be accessed through narrow tubes minimizing interaction with the outside world. The temperature difference between the two cells is measured by a thermoelectric device. When ligand is added to protein, the binding reaction occurs in the sample cell, and heat is generated or absorbed according to the thermodynamics of the reaction. The temperature difference between the sample and reference cell is kept constant at 0 by addition or removal of heat to the sample cell. The integral of the power required to maintain a constant temperature difference $\Delta T=0$ between the reference and sample cell, is a measure of total heat resulting from the binding reaction.

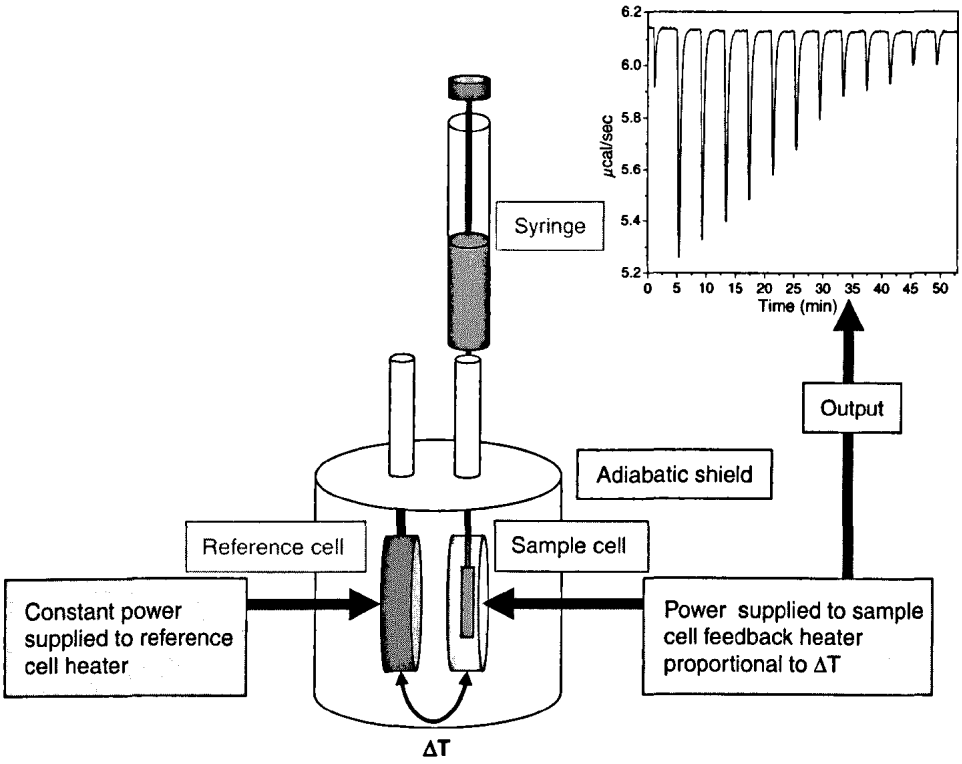


Figure 111: Schematic representation of a typical power consumption ITC. The diagram shows an oversimplification of how the power applied by the instrument to maintain constant temperature between the reference and sample cell is measured resulting in the instrument signal. Reprinted with permission from (*171*)**. Copyright 2008 Elsevier.**

If the concentrations of protein and ligand are appropriate, the first peaks will be of similar size, corresponding to full binding of all the ligand added. As saturation is reached, less and less ligand will bind and the peaks will decrease in size. The last peaks will represent the heat of dilution, as all protein has ligand bound. The integral of the peaks per mole of ligand plotted against the molar ratio of ligand to protein, gives a complete binding isotherm. The binding affinity is determined by the slope at which the peaks fall when saturation is reached (see Figure 112). The enthalpy of the reaction is the total difference between the first and last integrated peaks, and the stoichiometric number is found at the inflection point.
The advantages of ITC as a technique are that both ligand and protein are free in solution and do not have to be labeled or immobilized on a surface, as with e.g. surface plasmon resonance. Also, there are almost no buffer restrictions. The disadvantage is that high protein concentrations and significant amounts of protein are usually required to get reliable results, even through the protein consumption has been significantly lowered in the new equipment. Also, the vigorous stirring can cause precipitation of the protein if it is unstable. Furthermore, the use of reducing agents will cause some disturbance of the baseline due to the constant reaction. The producer of the ITC instrument (MicroCal) recommends the use of tris(2-carboxyethyl)phosphine (TCEP) or β-mercaptoethanol over DTT(*172*).

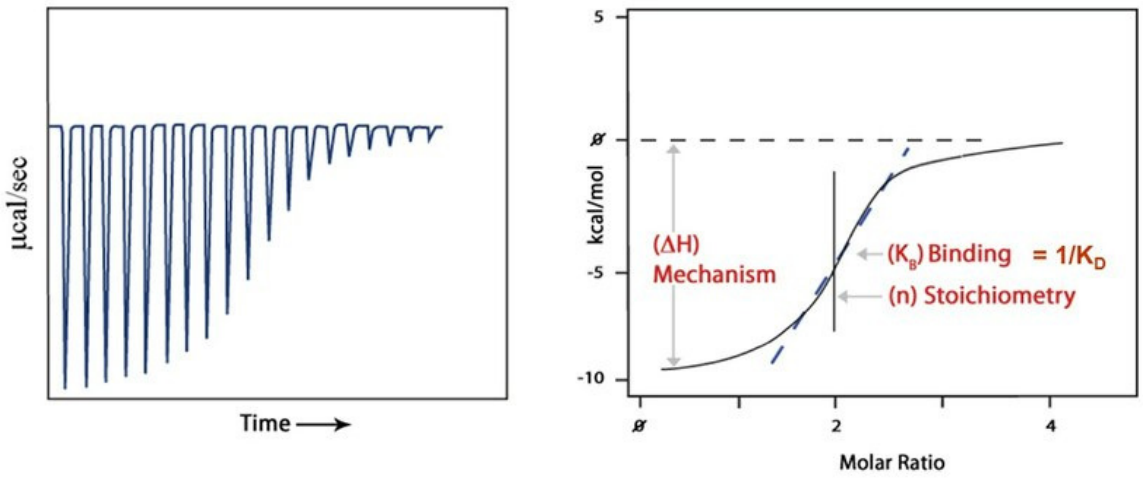


Figure 112: Typical ITC raw data (left) and integrated data (right) showing the binding isotherm and how the different parameters are determined. Reprinted with permission from (*173*)**. Copyright 2000-2009 GE Healthcare.**

6.5.1 Experimental procedures

In this section the experimental procedures for the isothermal titration calorimetry experiments will be described. The experiments were performed at the Faculty of Pharmaceutical Sciences at Copenhagen University under supervision of Associate Professor Lars Olsen.

The calorimetry experiments were performed using a microcalorimeter (ITC200, Microcal, MA, USA). The software ORIGIN 7.0 (Microcal, MA, USA) was used to determine the thermodynamic properties of ligand binding using non-linear least-squares fitting assuming a single-site model. For all experiments, the first peak was omitted from data analysis.

6.5.1.1 Preparation of protein samples

c*h*TPH2 purified as described in section 4.2.4.1 was thawed and mixed. The total volume was approx. 3.2 ml. The solution was added to a "Slide-A-Lyzer" cassette from Pierce (3-12 ml) with a molecular cut-off of 10000 Da, after the cassette had been hydrated for 2 min. in buffer. The buffer used for all experiments weas: 100 mM HEPES/NaOH, 100 mM $(NH_4)_2SO_4$, pH 7.00. The buffer used resembles that from the last step in the purification. The main differences are a higher buffer concentration, which should keep pH more stable when ligands are added, and a lower pH which is used in order to be able to compare the K_D values with K_M values determined previously using the kinetic assay(*13*).

The dialyze-cassette was placed in 1 L of buffer with stirring at 4 °C for 2 hours. The dialysis buffer was discarded and 800 ml new buffer was added. The dialysis setup was left overnight. The dialyzed protein solution was removed from the cassette and filtered (there was visible precipitation). The protein concentration was measured by UV-Vis at the NanoDrop 1000 spectrophotometer using ε_{280nm} = 39310 M^{-1} cm^{-1} to 225 µM (triple determination). The protein solution was frozen in aliquots of 350 µl in liquid nitrogen and kept at -80 °C until further use.

6.5.1.2 Prepration of ligand samples

The dialysis buffer from the second round of dialysis of the protein sample was used to make ligand solutions in order to ensure that the composition was exactly the same for protein and ligand samples. The ligands were weighed out and solubilized in buffer flushed with argon for 15 min to ensure a minimum of oxygen (BH_4 is sensitive to oxidation). After solubilization, pH was adjusted to 7.00 if necessary, and the

final solution was flushed with argon for 10 min. The concentrations of tryptophan and BH_2 were determined with UV-Vis at the NanoDrop 1000 spectrophotometer using the extinction coefficients shown in Table 29. The BH_4 sample was measured, then added 1 mM TCEP and 2 mM $(NH_4)_2Fe(II)(SO_4)_2$ to simulate the reducing assay conditions, and subsequently measured again. As the concentration of the BH_4 solution determined by UV-Vis did not correspond with the mass of BH_4 solubilized, the concentration calculated from the mass is given in Table 29.

Table 29: The concentration and extinction coefficients for ligand solutions for ITC experiments.

Ligand	Concentration	Extinction coefficient
Tryptophan	1.85 mM	ε_{278nm} = 5500 cm^{-1} M^{-1}(*174*)
BH_2	1.72 mM	ε_{330nm} = 6170 cm^{-1} M^{-1}(*175*)
BH_4	2 mM	-

6.5.1.3 ITC experiments

All experiments were performed at 15 °C, to ensure that the results can be compared to those from the kinetic assay, which is also performed at 15 °C.

At the beginning of each day and between different samples, the sample cell was cleaned several times with water using the cleaning program in the software and a manual Hamilton syringe. Prior to addition of protein solution, the sample cell was washed with buffer. The syringe was washed by filling with water and emptying at least 3 times between different samples. Furthermore, the syringe was washed with one volume of the ligand solution before use. The reference cell was filled with water and the sample cell was filled with buffer, water or protein using a Hamilton syringe. Excess solution was removed from the top of the sample cell, to ensure a correct volume of 206 µL.

Most experiments were run with 20 injections: the first of 0.5 µl to ensure that no air bobbles were trapped in the syringe and the following of each 2.0 µl. Generally, the time difference between each injection was 180 s, however, this was occasionally changed. The stirring speed was 1000 rpm and the differential power (DP) was set to 6 µcal/s. These are standard running conditions.

6.5.2 Results and discussion

In this section results from the ITC experiments will be presented and discussed.

6.5.2.1 Blanks

All ligand solutions were added to the buffer without any protein present to check how big peaks arise from differences in the two samples and from the mixing. The results for tryptophan and BH_2 can be seen in Figure 113.

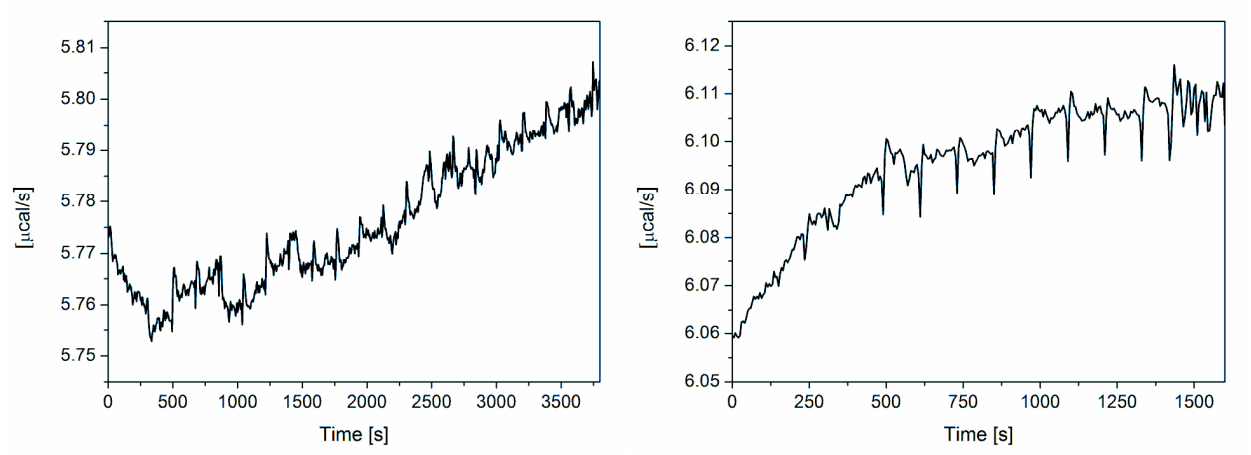


Figure 113: Left: Tryptophan injected in buffer, both with 1 mM TCEP. The temperature was 15 °C and 2 ul injections (first 0.5 µl) were made every 180 s. Right: BH_2 injected in buffer, both with 1 mM TCEP and 0.4 mM $(NH_4)_2Fe(II)(SO_4)_2$. The temperature was 15 °C and 2 µl injections (first 0.5 µl) were made every 180 s.

The small and almost indistinguishable peaks compared to the noise on the baseline for both tryptophan and BH_2 indicates that the ligand and buffer solutions are very much alike, and that the heat from mixing is small.

For BH_4, the picture is somewhat different, see Figure 114. The large decrease in the baseline following each injection suggests that a reaction is taking place. This also corresponds with the fact that the decreases get smaller and smaller for each injection, which is typical for reactions(*171*). BH_4 is prone to auto-oxidation by O_2(*176*), especially in non-acidic solutions, where there is no protonation of N5(*177*).

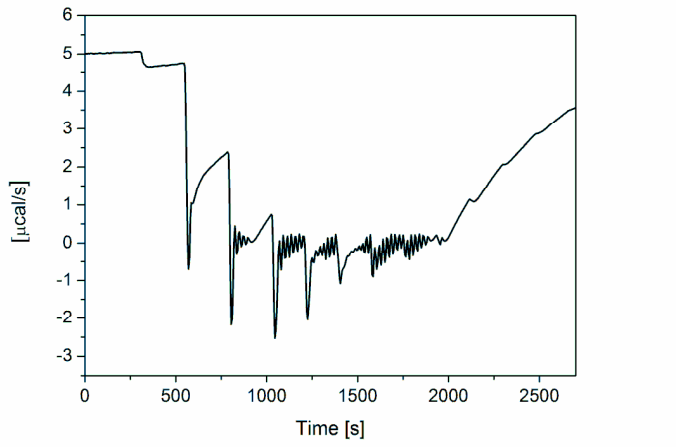


Figure 114: BH_4 injected in buffer, both with 1 mM TCEP and 0.4 mM $(NH_4)_2Fe(II)(SO_4)_2$. The temperature was 15 °C and 2 ul injections (first 0.5 µl) were made every 240 s.

The absorption spectrum of the BH_4 solution suggests that the main content is not BH_4, see Figure 115. Compared to the absorption spectrum of BH_4 in acidic solution, the shoulder at 266 nm typical for BH_4 is missing. Even upon reduction by addition of TCEP the shoulder does not appear. The spectrum of the BH_4 solution made in this project more resembles the spectrum of quinonoid dihydropterin q-BH_2, which is the product of the auto-oxidation of BH_4(*176*).

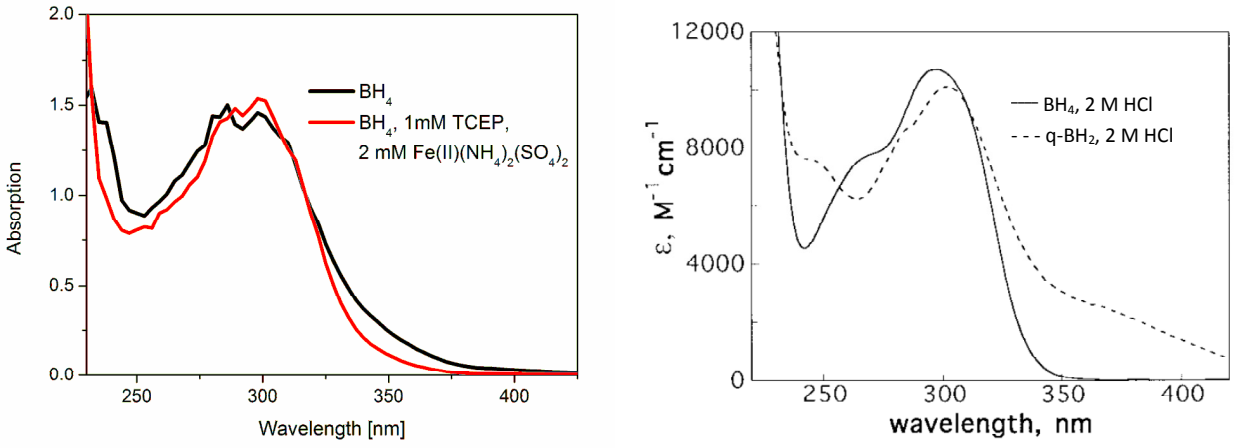


Figure 115: Left: Absorption spectrum of BH_4 before and after addition of 1 mM TCEP and 2 mM $(NH_4)_2Fe(II)(SO_4)_2$. Right: Absorption spectrum of BH_4 and quinonoid dihydropterin (q-BH_2) in 2 M HCl. Adapted with permission from (*174*). Copyright 1999 Elsevier.

One possible explanation for the pattern seen in Figure 114 is the reduction of q-BH_2 back to BH_4, when the ligand solution is added to the buffer solution containing 1 mM TCEP. Whether this is the case or not, it can be concluded that the BH_4 solution is not suitable for addition to protein samples, as the large reaction heats would disguise the binding heats. Furthermore, if the solution does contain q-BH_2 instead of BH_4 or a mixture of the two, the binding constants would be ambiguous. A solution to this problem could be to solubilize BH_4 in acidic solution, but as ITC requires that the protein and ligand is in the same solution this cannot be performed, as c*h*TPH2 is not stable at low pH, see section 4.3.3.3. Alternatively, the ITC experiment can be carried out under oxygen free conditions as in e.g. a glovebox.

6.5.2.2 Determination of binding constants

First, the binding constant of tryptophan to c*h*TPH2 was determined in a solution with 1 mM TCEP. This means that no Fe(II) was added, and thus the protein might not have been saturated with iron. After the measurement, protein precipitation was observed, which means that the actual concentration will also be lower than the 225 µM determined before the measurement. The actual protein concentration during the experiment is thus hard to determine. For this particular experiment only, TCEP was added from a very concentrated solution which caused precipitation of the protein even before addition to the sample cell. A lower protein concentration must be anticipated than in the other ITC experiments described here, where the actual protein concentration was determined by UV-Vis after each ITC experiment.

The protein precipitation, and following change in active protein concentration, should not affect the determination of the association constant for optimally designed experiments. However, if the precipitation process takes place during measurements, the baseline might be affected due to the constant release of heat from the phase change. After addition of protein to the sample cell, the stirring was started and after only 5 min, the protein solution was checked and precipitation was observed. This indicates that the fast stirring (1000 rpm) might be responsible for the precipitation. This hopefully means that most protein precipitation happens during temperature and baseline equilibration, and thus should not affect the actual measurement. For the rest of the samples, the protein solution was centrifuged after each measurement, and the actual protein concentration was determined by UV-Vis. Precipitation of protein due to stirring is a well-known problem when performing ITC.

The baseline-corrected data for the ITC experiment of tryptophan binding to c*h*TPH2 is seen in Figure 116, left, and displays a clear binding of tryptophan to the protein. The binding heat is negative (positive peaks), which means that the binding reaction of tryptophan to the protein is endothermic, and thus must be entropically driven. This corresponds well with the fact that tryptophan binds in a hydrophobic pocket. The design of the experiment is not optimal, as the protein concentration is too low (or the ligand concentration too high) and thus no plateau is seen for the integrated data of the first peaks. This makes determination of the dissociation constant difficult. Furthermore, the protein concentration was not measured after the experiment, so instead it was estimated. In order to fit the data, the stoichiometry parameter N had to be fixed to 1. The association constants determined from the fitting was dependent on protein concentration and varied between $9.987 \cdot 10^3$ M^{-1} (1 µM protein) and $1.749 \cdot 10^4$ M^{-1} (50 µM protein), corresponding to dissociation constants of 100 and 57 µM, respectively. These dissociation constants are within the same order of magnitude as the kinetically determined K_M value for tryptophan, which is 15 µM(*13*). In order to determine an exact dissociation constant by ITC, the experiment will have to be repeated with a higher protein concentration or lower tryptophan concentration, so the data integration can be performed independent of protein concentration.

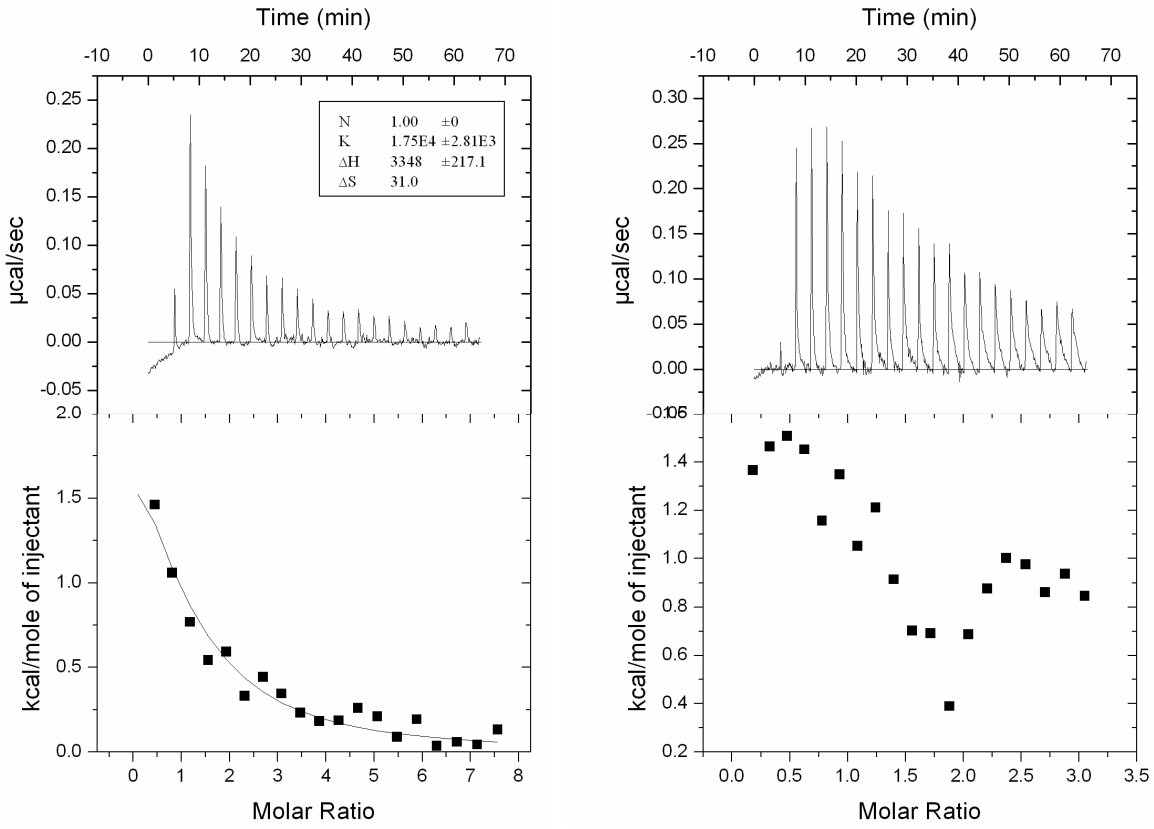


Figure 116: ITC experiments with tryptophan binding. The top graph shows the baseline-corrected raw data, and the bottom graph the integrated data along with the fit. Left: Tryptophan injected in c*h*TPH2 solution, both with 1 mM TCEP. The temperature was 15 °C and 2 µl injections (first: 0.5 µl) were made every 180 s. The data was fitted assuming a protein concentration of 50 µM and N fixed to 1, which gave a dissociation constant of 57 µM. Right: Tryptophan injected in c*h*TPH2+BH_2 solution, both with 1 mM TCEP and 0.4 mM $(NH_4)_2Fe(II)(SO_4)_2$.. The temperature was 15 °C and 2 µl injections (first: 0.5 µl) were made every 180 s. No fitting could be made due to the widening of peaks at the end of the run and a steep change in baseline.

To check if the binding of substrate and cofactor is independent, the binding of tryptophan to the protein with BH_2 already bound was also investigated. As tryptophan and BH_2 bind in separate binding pockets

according to the crystal structures(*12*, *47*) this is not expected, but substrate binding induced conformational changes in the protein, could change the association constant. The result can be seen in Figure 116, right, where the general pattern suggests binding of tryptophan. Fitting of the data was not possible, due to a steep change in baseline and widening of the last peaks which made integration unreliable. The widening of the last peaks is possibly caused by air in the system. However, it can be concluded that tryptophan does bind to the protein with bound BH_2, and the binding is endothermic.

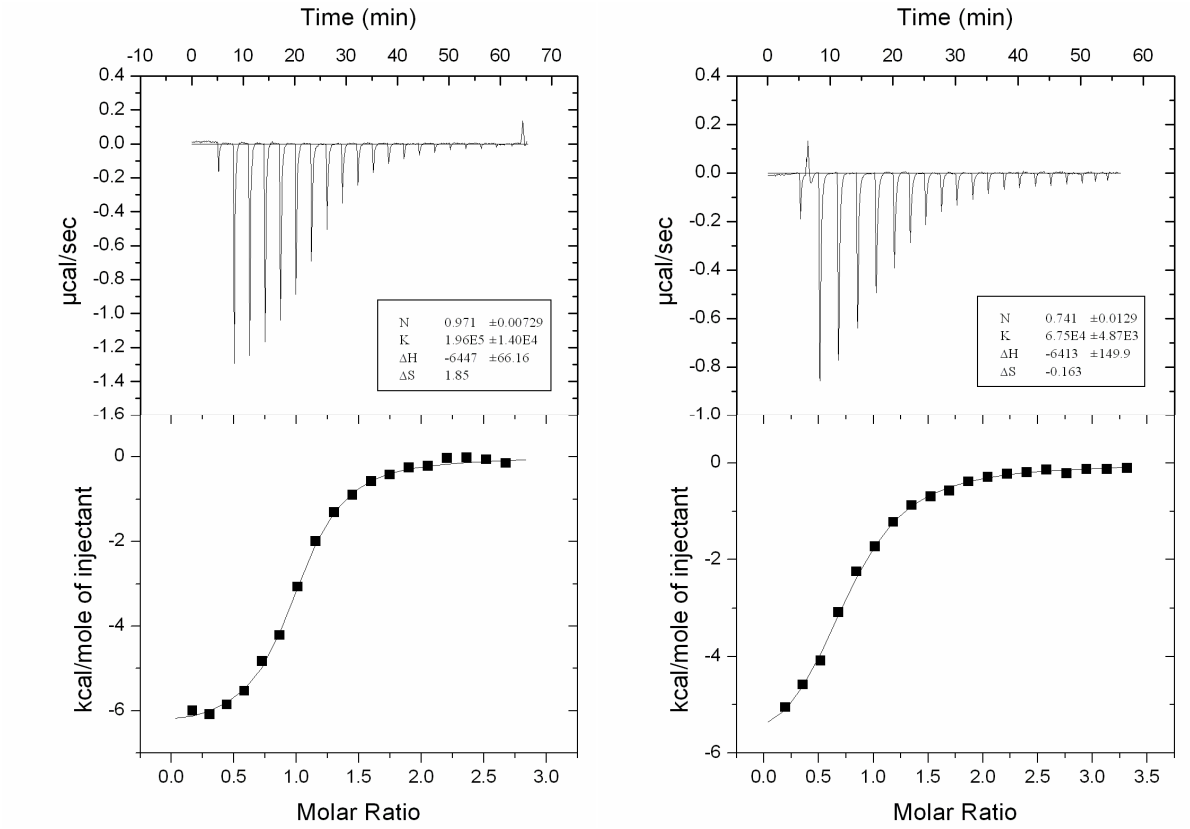


Figure 117: ITC experiments with BH_2 binding. The top graph shows the baseline-corrected raw data, and the bottom graph the integrated data along with the fit. Left: BH_2 injected in *ch*TPH2 solution, both with 1 mM TCEP. The temperature was 15 °C and 2 µl injections (first: 0.5 µl) were made every 180 s. The dissociations constant was determined to 6 µM for a protein concentration of 124 µM. Right: BH_2 injected in *ch*TPH2+tryptophan solution, both with 1 mM TCEP and 0.4 mM $(NH_4)_2Fe(II)(SO_4)_2$. The temperature was 15 °C and 2 µl injections (first 0.5 µl) were made every 180 s. The dissociations constant was determined to 15 µM for a protein concentration of 106 µM.

BH_2 was injected directly into a protein sample with no other ligands present. The baseline-corrected raw data and the integrated data can be seen in Figure 117, left. Fitting of the integrated data gave an association constant of $1.956 \cdot 10^5$ M^{-1}, which corresponds to a dissociation constant of 5 µM. For the protein concentration measured after the experiment (124 µM) N=0.97, which is very close to 1. It can thus be concluded that BH_2 binds 1:1 with the protein, and that the reaction is exothermic (negative peaks). The binding of BH_2 is significantly stronger than binding of tryptophan corresponding with the smaller dissociation constant, but this was also expected since BH_2 is an inhibitor of TPH. Fitting of the integrated data could be performed without fixing any parameters, suggesting that the protein and BH_2 concentrations are suitable.

BH_2 was also injected into a sample containing c*h*TPH2 and tryptophan. The result is seen in Figure 117. right. The integrated data gives an association constant of $6.747 \cdot 10^4$ M^{-1}, which equals a dissociation constant of 15 µM. The protein concentration after this measurement and centrifugation was 106 µM, which gives N=0.74, corresponding with a 1:1 binding. The difference in the dissociation constant between the protein alone (5 µM), and the protein with bound tryptophan (15 µM) is too small to be significant, especially since the experiments were only performed once. More reliable data for BH_2 binding to c*h*TPH2 with already bound tryptophan could be achieved by lowering the ligand concentration or increasing the protein concentration.

In order to determine whether c*h*TPH1 and c*h*TPH2 differs significantly in binding affinity for the cofactor as suggested by the kinetic data, ITC measurements of TPH1 must be carried out. The recent development of purification procedures for c*h*TPH1 and *h*TPH1 makes this possible, as both can be concentrated to around 200 µM, which should be sufficient for ITC measurements. As the natural cofactor cannot be used for ITC experiments, the dissociation constants for BH_2 or other cofactor analogues could be compared.

6.5.3 Conclusions

In general, protein precipitation made the ITC measurements difficult. Despite this several results were obtained. It can be concluded that both tryptophan and the inhibitor BH_2 binds to c*h*TPH2. The dissociation constant of BH_2 is 5 µM, while the dissociation constant for tryptophan is somewhat higher and within the same order of magnitude as K_M determined by a kinetic assay. The precise value cannot be determined from the experiments performed here. Furthermore, tryptophan and BH_2 binds independently of each other, corresponding with different binding pockets also seen in crystal structures. Tryptophan binding is entropically driven, while binding of BH_2 is exothermic.

6.6 Conclusions

In this chapter, c*h*TPH2 has been investigated using several different techniques in order to provide information on the reaction mechanism.

Four mutants of c*h*TPH2 were made, and it could be concluded that Glu363, which is an iron ligand, is essential for activity, while Tyr358 affects the activity, but is not absolutely required. Both amino acids were important for the solubility of the protein. CD and MCD measurements were performed on c*h*TPH2 samples added either tryptophan, BH_4 or both. The results clearly demonstrated that the active site iron changes conformation from 6C to 5C only when both substrate and cofactor binds. A difference between samples made in glycerol and sucrose was observed, and this effect was tested further by absorption measurements. It was clear that addition of glycerol to c*h*TPH2 with bound Fe(III) gave a yellow color indicative of interaction. DSF measurements demonstrated that c*h*TPH2 and ct*h*TPH2 is significantly more stable with Fe(II) bound than with Fe(III). Furthermore, the addition of substrates and cofactor only stabilizes c*h*TPH2 when Fe(II) is present. ITC measurements were performed on c*h*TPH2. Despite problems with protein precipitation, the dissociations constant of BH_2 was determined to 5 µM, while the dissociation constant for tryptophan is somewhat higher. Tryptophan and BH_2 bind independently of each other.

7 Overall conclusions

The two overall goals of this project were: I) To develop purification methods for full-length human TPH1 and TPH2, and II) to characterize purified variants of TPH to achieve a better understanding of the structure-functional relationship.

A successful purification method for full-length *h*TPH1 was developed, which resulted in pure, active and stable protein. The method includes affinity-purification using MBP tag and high salt concentrations in all buffers. It is the first time a successful purification method for *h*TPH1 has been presented, which also yields stable protein after cleavage from the affinity tag. This allows for a thorough characterization of the full-length human enzyme, which has not been possible before. Furthermore, successful purification procedures were developed for a number of other TPH variants. A purification method for rc*h*TPH1 similar to that of *h*TPH1 using MBP affinity tag was developed, and revealed that the regulatory domain causes dimerization. Also, a new purification method for c*h*TPH1 using GST affinity-tag was developed, which eliminated impurities identified by mass spectrometry in earlier projects. rc*h*TPH2 and *h*TPH2 were attempted purified using the same general method as for *h*TPH1, but the yields were very low and the products were unstable and could not be concentrated. The high salt concentration does thus not seem to stabilize the regulatory domain of TPH2. *h*TPH1 displayed a significantly higher specific activity than c*h*TPH1 and rc*h*TPH1, suggesting that the tetramerization increase reactivity, possibly by cooperativity. The stability of c*h*TPH1, *h*TPH1, c*h*TPH2 and ct*h*TPH2 were studied by differential scanning fluorimetry, and the results gave information on the effect of pH, salt type and ionic strength.

The results from crystallization were unfortunately rather limited. The crystallization procedure for c*g*TPH1 was optimized to faster crystal growth by addition of tryptophan and incubation at room temperature. Crystals without imidazole in the crystallization conditions could be obtained. The solved structures were however of poor quality and did not contribute with any new information on TPH. A strategic approach to crystallization of TPH variants was developed, and the 5 suitable TPH variants were purified followed by immediate screen set-up eliminating the freezing-thawing step. No crystals of any TPH2 variant were found. However, the pipeline screening produced leads for rc*h*TPH1 and several crystals for c*h*TPH1.

In the mechanistic studies, several different techniques were used to investigate TPH. Four mutants of c*h*TPH2 were made, and it could be concluded that Glu363, which is an iron ligand, is essential for activity, while Tyr358 affects the activity, but is not absolutely required. CD and MCD results on c*h*TPH2 clearly demonstrated that the active site iron changes conformation from 6C to 5C only when both substrate and cofactor binds. It is the first time a MCD study has been performed any AAAH with the natural cofactor, and the quality of the data was very high. A difference between samples made in glycerol and sucrose was observed, and this effect was tested further by absorption measurements. It was clear that addition of glycerol to c*h*TPH2 with bound Fe(III) gave a yellow color indicative of interaction. DSF measurements demonstrated that c*h*TPH2 and ct*h*TPH2 are significantly more stable in the Fe(II) form compared to the Fe(III) form. Furthermore, the addition of substrates and cofactor only stabilizes c*h*TPH2 when Fe(II) is present. ITC measurements were performed on c*h*TPH2. Despite problems with protein precipitation, the

dissociation constant of BH_2 was determined to 5 µM, while the dissociation constant for tryptophan is somewhat higher. Tryptophan and BH_2 bind independently of each other.
Overall, the work performed in this project has made significant contributions to the research on TPH, both by providing new purification procedures as well as advancements in the knowledge on the mechanism of TPH.

7.1 Future work

The results presented in this thesis provide a good starting point for further research on TPH. The purification procedures involving affinity tags could be shortened significantly by introducing on-column cleavage. For purifications using the MBP column, a thorough investigation of the first affinity step should be performed in order to elucidate why a large portion of the MBP-tagged protein is found in the flow-through and cannot subsequently bind to the column. Kinetic characterization of the purified TPH variants should also be made. The availability of pure and stable TPH variants containing the regulatory domain can be used for investigation of the regulatory mechanisms including possible feed-back regulation by e.g. serotonin or inhibition by substrates.
Regarding crystallization, more crystals of c*g*TPH1 without imidazole in the crystallization conditions should be produced and tested for diffraction in the hope of achieving high resolution data. Furthermore, the leads for crystallization of c*h*TPH1 should be pursued in order to solve the structure without any substrates bound, or with either tryptophan or tryptophan and BH_2 bound.
The ITC experiments with c*h*TPH2 performed here were only preliminary, but provide information on how to design suitable experiments for determination of dissociation constants for tryptophan and BH_2. Such experiments should be performed, and if possible compared to similar ITC experiments of TPH1 to determine any differences between the isoforms. The ITC experiments could also be performed with unnatural but redox-active cofactor analogues such as 6-methyltetrahydropterin for more reliable results as compared to BH_4.

7.2 Outlook

It is the hope that the continued work on TPH regarding purification and characterization in time will lead to a thorough understanding of the structure, mechanism and regulation as well as shed further light on the differences between isoform 1 and 2. Even though the work is inhibited by the inherent instability of TPH, small steps of progress are made with each project. In the future, knowledge obtained from the research on TPH might aid in developing improved treatments for serotonin-related disorders.

8 References

1. Lovenberg, W., Jequier, E., and Sjoerdsma, A. (1967) Tryptophan hydroxylation: measurement in pineal gland, brainstem, and carcinoid tumor, *Science 155*, 217-219.

2. Mockus, S. M., and Vrana, K. E. (1998) Advances in the molecular characterization of tryptophan hydroxylase, *Journal of Molecular Neuroscience 10*, 163-179.

3. Ganguly, S., Coon, S. L., and Klein, D. C. (2002) Control of melatonin synthesis in the mammalian pineal gland: the critical role of serotonin acetylation, *Cell and Tissue Research 309*, 127-137.

4. Lucki, I. (1998) The spectrum of behaviors influenced by serotonin, *Biological Psychiatry 44*, 151-162.

5. Martínez, A., Knappskog, P. M., and Haavik, J. (2001) A structural approach into human tryptophan hydroxylase and its implications for the regulation of serotonin biosynthesis, *Current Medicinal Chemistry 8*, 1077-1091.

6. Walther, D. J., Peter, J.-U., Bashammakh, S., Hörtnagl, H., Voits, M., Fink, H., and Bader, M. (2003) Synthesis of serotonin by a second tryptophan hydroxylase isoform, *Science 299*, 76.

7. Walther, D. J., and Bader, M. (2003) A unique central tryptophan hydroxylase isoform, *Biochemical Pharmacology 66*, 1673-1680.

8. Carkaci-Salli, N., Flanagan, J. M., Martz, M. K., Salli, U., Walther, D. J., Bader, M., and Vrana, K. E. (2006) Functional domains of human tryptophan hydroxylase 2 (hTPH2), *The Journal of Biological Chemistry 281*, 28105-28112.

9. Nielsen, M. S. (2007) Expression, purification and characterization of tryptophan hydroxylase, Ph.D. thesis. Department of Chemistry, Technical University of Denmark.

10. Boesen, J. (2010) Cloning, expression, purification and characterization of tryptophan hydroxylase variants, Ph.D. thesis. Department of Chemistry, Technical University of Denmark.

11. Haahr, L. T. (2008) Expression, purification and structural studies of tetrameric variants of tryptophan hydroxylase isoform 2, Master thesis. Department of Chemistry, Technical University of Denmark.

12. Windahl, M. S., Petersen, C. R., Christensen, H. E. M., and Harris, P. (2008) Crystal structure of tryptophan hydroxylase with bound amino acid substrate, *Biochemistry 47*, 12087-12094.

13. Windahl, M. S., Boesen, J., Karlsen, P. E., and Christensen, H. E. (2009) Expression, purification and enzymatic characterization of the catalytic domains of human tryptophan hydroxylase isoforms, *The Protein Journal 28*, 400-406.

14. Fitzpatrick, P. F. (2003) Mechanism of aromatic amino acid hydroxylation, *Biochemistry 41*, 14083-14091.

15. Thöny, B., Auerbach, G., and Blau, N. (2000) Tetrahydrobiopterin biosynthesis, regeneration and functions, *The Biochemical Journal 347 Pt 1*, 1-16.

16. Fitzpatrick, P. F. (1999) Tetrahydrobiopterin-dependent amino acid hydroxylases, *Annual Review of Biochemistry 68*, 355-381.

17. Fitzpatrick, P. F. (2000) The aromatic amino acid hydroxylases, *Advances in Enzymology and Related Areas of Molecular Biology 74*, 235-294.

18. Kaufman, S. (1957) The enzymatic conversion of phenylalanine to tyrosine, *Journal of Biological Chemistry 226*, 511-524.

19. Kaufman, S. (1993) The phenylalanine hydroxylating system, *Advances in Enzymology and Related Areas of Molecular Biology 67*, 77-264.

20. Flatmark, T., and Stevens, R. C. (1999) Structural insight into the aromatic amino acid hydroxylases and their disease-related mutant forms, *Chemical Reviews 99*, 2137-2160.

21. Fenylketonuri (PKU, Føllings sygdom) - Statens Serum Institut. http://www.ssi.dk/Diagnostik/Klinisk%20information/Blodproeve%20fra%20nyfoedte/Sygdomme%20som%20indgaer%20i%20screeningen/Fenylketonuri.aspx, 27/4-2012.

22. Nagatsu, T., Levitt, M., and Udenfriend, S. (1964) Tyrosine hydroxylase - the initial step in norepinephrine biosynthesis, *The Journal of Biological Chemistry 239*, 2910-2917.

23. Knappskog, P., Flatmark, T., Mallet, J., Luedecke, B., and Bartholome, K. (1995) Recessively inherited L-DOPA-responsive dystonia caused by a point mutation (Q381K) in the tyrosine hydroxylase gene, *Human Molecular Genetics 4*, 1209-1212.

24. Bademci, G., Edwards, T. L., Torres, A. L., Scott, W. K., Züchner, S., Martin, E. R., Vance, J. M., and Wang, L. (2010) A rare novel deletion of the tyrosine hydroxylase gene in Parkinson disease, *Human Mutation 31*, E1767-E1771.

25. Lüdecke, B., Knappskog, P. M., Clayton, P. T., Surtees, R. A., Clelland, J. D., Heales, S. J., Brand, M. P., Bartholomé, K., and Flatmark, T. (1996) Recessively inherited L-DOPA-responsive parkinsonism in infancy caused by a point mutation (L205P) in the tyrosine hydroxylase gene, *Human Molecular Genetics 5*, 1023-1028.

26. Pavon, J. A., and Fitzpatrick, P. F. (2006) Insights into the catalytic mechanisms of phenylalanine and tryptophan hydroxylase from kinetic isotope effects on aromatic hydroxylation, *Biochemistry 45*, 11030-11037.

27. Slominski, A., Pisarchik, A., Johansson, O., Jing, C., Semak, I., Slugocki, G., and Wortsman, J. (2003) Tryptophan hydroxylase expression in human skin cells, *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease 1639*, 80-86.

28. Patel, P. D., Pontrello, C., and Burke, S. (2004) Robust and tissue-specific expression of TPH2 versus TPH1 in rat raphe and pineal gland, *Biological Psychiatry 55*, 428-433.

29. Zill, P., Büttner, A., Eisenmenger, W., Möller, H.-J., Ackenheil, M., and Bondy, B. (2007) Analysis of tryptophan hydroxylase I and II mRNA expression in the human brain: a post-mortem study, *Journal of Psychiatric Research 41*, 168-173.

30. Hasegawa, H., Yanagisawa, M., Inoue, F., Yanaihara, N., and Ichiyama, A. (1987) Demonstration of non-neural tryptophan 5-mono-oxygenase in mouse intestinal mucosa, *The Biochemical Journal 248*, 501-509.

31. Zhang, X., Beaulieu, J.-M., Sotnikova, T. D., Gainetdinov, R. R., and Caron, M. G. (2004) Tryptophan hydroxylase-2 controls brain serotonin synthesis, *Science 305*, 217.

32. Gutknecht, L., Waider, J., Kraft, S., Kriegebaum, C., Holtmann, B., Reif, A., Schmitt, A., and Lesch, K.-P. (2008) Deficiency of brain 5-HT synthesis but serotonergic neuron formation in Tph2 knockout mice., *Journal of Neural Transmission 115*, 1127-1132.

33. Murphy, K. L., Zhang, X., Gainetdinov, R. R., Beaulieu, J.-M., and Caron, M. G. (2008) A regulatory domain in the N terminus of tryptophan hydroxylase 2 controls enzyme expression, *The Journal of Biological Chemistry 283*, 13216-13224.

34. Tenner, K., Walther, D., and Bader, M. (2007) Influence of human tryptophan hydroxylase 2 N- and C-terminus on enzymatic activity and oligomerization, *Journal of Neurochemistry 102*, 1887-1894.

35. Camilleri, M. (2011) LX-1031, a tryptophan 5-hydroxylase inhibitor, and its potential in chronic diarrhea associated with increased serotonin, *Neurogastroenterology and Motility 23*, 193-200.

36. Liu, Q., Yang, Q., Sun, W., Vogel, P., Heydorn, W., Yu, X.-qing, Hu, Z., Yu, W., Jonas, B., Pineda, R., Calderon-gay, V., Germann, M., Neill, E. O., Brommage, R., Cullinan, E., Platt, K., Wilson, A., Powell, D., Sands, A., Zambrowicz, B., and Shi, Z.-cai. (2008) Discovery and characterization of novel tryptophan hydroxylase inhibitors that selectively inhibit serotonin synthesis in the gastrointestinal tract, *The journal of Pharmacology and Experimental Therapeutics 325*, 47-55.

37. Margolis, K. G., Stevanovic, K. D., Yang, Q. M., Li, Z., Mazo, R., and Gershon, M. D. (2011) An inhibitor of tryptophan hydroxylase successfully ameliorates TNBS-induced colitis, *Gastroenterology 140*, S-478.

38. Walitza, S., Renner, T. J., Dempfle, A., Konrad, K., Wewetzer, C., Halbach, A., Herpertz-Dahlmann, B., Remschmidt, H., Smidt, J., Linder, M., Flierl, L., Knölker, U., Friedel, S., Schäfer, H., Gross, C., Hebebrand, J., Warnke, A., and Lesch, K. P. (2005) Transmission disequilibrium of polymorphic variants in the tryptophan hydroxylase-2 gene in attention-deficit/hyperactivity disorder, *Molecular Psychiatry 10*, 1126-1132.

39. Cichon, S., Winge, I., Mattheisen, M., Georgi, A., Karpushova, A., Freudenberg, J., Freudenberg-Hua, Y., Babadjanova, G., Van Den Bogaert, A., Abramova, L. I., Kapiletti, S., Knappskog, P. M., McKinney, J., Maier, W., Jamra, R. A., Schulze, T. G., Schumacher, J., Propping, P., Rietschel, M., Haavik, J., and Nöthen, M. M. (2008) Brain-specific tryptophan hydroxylase 2 (TPH2): A functional Pro206Ser substitution and variation in the 5'-region are associated with bipolar affective disorder, *Human Molecular Genetics 17*, 87-97.

40. Zhang, X., Gainetdinov, R. R., Beaulieu, J.-M., Sotnikova, T. D., Burch, L. H., Williams, R. B., Schwartz, D. a, Krishnan, K. R. R., and Caron, M. G. (2005) Loss-of-function mutation in tryptophan hydroxylase-2 identified in unipolar major depression, *Neuron 45*, 11-16.

41. Zhou, Z., Roy, A., Lipsky, R., Kuchipudi, K., Zhu, G., Taubman, J., Enoch, M.-A., Virkkunen, M., and Goldman, D. (2005) Haplotype-based linkage of tryptophan hydroxylase 2 to suicide attempt, major depression, and cerebrospinal fluid 5-hydroxyindoleacetic acid in 4 populations, *Archives of General Psychiatry 62*, 1109-1118.

42. Van den Bogaert, A., Sleegers, K., De Zutter, S., Heyrman, L., Norrback, K.-F., Adolfsson, R., Van Broeckhoven, C., and Del-Favero Jr., J. (2006) Association of brain-specific tryptophan hydroxylase, TPH2, with unipolar and bipolar disorder in a Northern Swedish, isolated population, *Archives of General Psychiatry 63*, 1103-1110.

43. Mössner, R., Walitza, S., Geller, F., Scherag, A., Gutknecht, L., Jacob, C., Bogusch, L., Remschmidt, H., Simons, M., Herpertz-Dahlmann, B., Fleischhaker, C., Schulz, E., Warnke, A., Hinney, A., Wewetzer, C., and Lesch, K.-P. (2006) Transmission disequilibrium of polymorphic variants in the tryptophan hydroxylase-2 gene in children and adolescents with obsessive-compulsive disorder, *The International Journal of Neuropsychopharmacology 9*, 437-442.

44. Sheehan, K., Lowe, N., Kirley, A., Mullins, C., Fitzgerald, M., Gill, M., and Hawi, Z. (2005) Tryptophan hydroxylase 2 (TPH2) gene variants associated with ADHD, *Molecular Psychiatry 10*, 944-949.

45. Matthes, S., Mosienko, V., Bashammakh, S., Alenina, N., and Bader, M. (2010) Tryptophan hydroxylase as novel target for the treatment of depressive disorders, *Pharmacology 85*, 95-109.

46. Popova, N. K., and Kulikov, A. V. (2010) Targeting tryptophan hydroxylase 2 in affective disorder, *Expert Opinion on Therapeutic Targets 14*, 1259-1271.

47. Wang, L., Erlandsen, H., Haavik, J., Knappskog, P. M., and Stevens, R. C. (2002) Three-dimensional structure of human tryptophan hydroxylase and its implications for the biosynthesis of the neurotransmitters serotonin and melatonin, *Biochemistry 41*, 12569-12574.

48. Cianchetta, G., Stouch, T., Yu, W., Shi, Z. C., Tari, L. W., Swanson, R. V., Hunter, M. J., Hoffman, I. D., and Liu, Q. (2010) Mechanism of inhibition of novel tryptophan hydroxylase inhibitors revealed by co-crystal structures and kinetic analysis, *Current Chemical Genomics 4*, 19-26.

49. Kobe, B., Jennings, I. G., House, C. M., Michell, B. J., Goodwill, K. E., Santarsiero, B. D., Stevens, R. C., Cotton, R. G. H., and Kemp, B. E. (1999) Structural basis of autoregulation of phenylalanine hydroxylase, *Nature Structural Biology 6*, 442-448.

50. DeLano, W. L. (2002) The PyMOL Molecular Graphics System. http://www.pymol.org. Schrödinger LLC, San Carlos, CA, USA.

51. McKinney, J., Knappskog, P. M., and Haavik, J. (2005) Different properties of the central and peripheral forms of human tryptophan hydroxylase, *Journal of Neurochemistry 92*, 311-320.

52. Johansen, P. A., Jennings, I., Cotton, R. G., and Kuhn, D. M. (1995) Tryptophan hydroxylase is phosphorylated by protein kinase A, *Journal of Neurochemistry 65*, 882-888.

53. Winge, I., McKinney, J. A., Ying, M., D'Santos, C. S., Kleppe, R., Knappskog, P. M., and Haavik, J. (2008) Activation and stabilization of human tryptophan hydroxylase 2 by phosphorylation and 14-3-3 binding, *The Biochemical Journal 410*, 195-204.

54. Kuhn, D. M., Sakowski, S. A., Geddes, T. J., Wilkerson, C., and Haycock, J. W. (2007) Phosphorylation and activation of tryptophan hydroxylase 2: identification of serine-19 as the substrate site for calcium, calmodulin-dependent protein kinase II, *Journal of Neurochemistry 103*, 1567-1573.

55. Banik, U., Wang, G. A., Wagner, P. D., and Kaufman, S. (1997) Interaction of phosphorylated tryptophan hydroxylase with 14-3-3 proteins, *The Journal of Biological Chemistry 272*, 26219-26225.

56. Furukawa, Y., Ikuta, N., Omata, S., Yamauchi, T., Isobe, T., and Ichimura, T. (1993) Demonstration of the phosphorylation-dependent interaction of tryptophan hydroxylase with the 14-3-3 protein, *Biochemical and Biophysical Research Communications 194*, 144-149.

57. Chen, G.-L., and Miller, G. M. (2012) Advances in tryptophan hydroxylase-2 gene expression regulation: new insights into serotonin-stress interaction and clinical implications, *American Journal of Medical Genetics, Part B, Neuropsychiatric Genetics 159B*, 152-171.

58. Koehntop, K. D., Emerson, J. P., and Que, L. (2005) The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes, *Journal of Biological Inorganic Chemistry 10*, 87-93.

59. McKinney, J., Teigen, K., Frøystein, N. A., Salaün, C., Knappskog, P. M., Haavik, J., and Martínez, A. (2001) Conformation of the substrate and pterin cofactor bound to human tryptophan hydroxylase. Important role of Phe313 in substrate specificity, *Biochemistry 40*, 15591-15601.

60. Daubner, S. C., Moran, G. R., and Fitzpatrick, P. F. (2002) Role of tryptophan hydroxylase phe313 in determining substrate specificity, *Biochemical and Biophysical Research Communications 292*, 639-641.

61. Kino, K., Hara, R., and Nozawa, A. (2009) Enhancement of l-tryptophan 5-hydroxylation activity by structure-based modification of l-phenylalanine 4-hydroxylase from Chromobacterium violaceum, *Journal of Bioscience and Bioengineering 108*, 184-189.

62. Jiang, G. C., Yohrling, G. J., Schmitt, J. D., Vrana, K. E., and Schmitt, I. V. (2000) Identification of substrate orienting and phosphorylation sites within tryptophan hydroxylase using homology-based molecular modeling, *Journal of Molecular Biology 302*, 1005-1017.

63. Goodwill, K. E., Sabatier, C., Marks, C., Raag, R., Fitzpatrick, P. F., and Stevens, R. C. (1997) Crystal structure of tyrosine hydroxylase at 2.3 A and its implications for inherited neurodegenerative diseases, *Nature Structural Biology 4*, 578-585.

64. Fusetti, F., Erlandsen, H., Flatmark, T., and Stevens, R. C. (1998) Structure of tetrameric human phenylalanine hydroxylase and its implications for phenylketonuria, *Journal of Biological Chemistry 273*, 16962-16967.

65. Andersen, O. A., Flatmark, T., and Hough, E. (2001) High resolution crystal structures of the catalytic domain of human phenylalanine hydroxylase in its catalytically active Fe(II) form and binary complex with tetrahydrobiopterin, *Journal of Molecular Biology 314*, 279-291.

66. Andersen, O. A., Stokka, A. J., Flatmark, T., and Hough, E. (2003) 2.0 angstrom resolution crystal structures of the ternary complexes of human phenylalanine hydroxylase catalytic domain with tetrahydrobiopterin and 3-(2-thienyl)-L-alanine or L-norleucine: Substrate specificity and molecular motions related to substrate, *Journal of Molecular Biology 333*, 747-757.

67. Kemsley, J. N., Zaleski, K. L., Caradonna, J. P., and Solomon, E. I. (1999) Circular dichroism and magnetic circular dichroism spectrocscopy of the catalytically competent ferrous active site of phenylalanine hydroxylase and its interaction with pterin cofactor, *Journal of the American Chemical Society 121*, 1528-1536.

68. Wasinger, E. C., Mitic, N., Hedman, B., Caradonna, J., Solomon, E. I., and Hodgson, K. O. (2002) X-ray absorption spectroscopic investigation of the resting ferrous and cosubstrate-bound active sites of phenylalanine hydroxylase, *Biochemistry 41*, 6211-6217.

69. Loeb, K. E., Westre, T. E., Kappock, T. J., Glasfeld, E., Caradonna, J. P., Hedman, B., Hodgson, K. O., and Solomon, E. I. (1997) Spectroscopic characterization of the catalytically competent ferrous site of the resting , activated , and substrate-bound forms of phenylalanine hydroxylase, *Journal of the American Chemical Society 119*, 1901-1915.

70. Chow, M. S., Eser, B. E., Wilson, S. a, Hodgson, K. O., Hedman, B., Fitzpatrick, P. F., and Solomon, E. I. (2009) Spectroscopy and kinetics of wild-type and mutant tyrosine hydroxylase: mechanistic insight into O2 activation., *Journal of the American Chemical Society 131*, 7685-7698.

71. Fitzpatrick, P. F. (1991) Steady-state kinetic mechanism of rat tyrosine hydroxylase, *Biochemistry 30*, 3658-3662.

72. Pember, S. O., Johnson, K. A., Villafranca, J. J., and Benkovic, S. J. (1989) Mechanistic studies on phenylalanine hydroxylase from chromobacterium-violaceum. Evidence for the formation of an enzyme-oxygen complex, *Biochemistry 28*, 2124-2130.

73. Volner, A., Zoidakis, J., and Abu-Omar, M. M. (2003) Order of substrate binding in bacterial phenylalanine hydroxylase and its mechanistic implication for pterin-dependent oxygenases, *Journal of Biological Inorganic Chemistry 8*, 121-128.

74. Bassan, A., Blomberg, M. R. a, and Siegbahn, P. E. M. (2003) Mechanism of dioxygen cleavage in tetrahydrobiopterin-dependent amino acid hydroxylases, *Chemistry 9*, 106-115.

75. Bassan, A., Borowski, T., and Siegbahn, P. E. M. (2004) Quantum chemical studies of dioxygen activation by mononuclear non-heme iron enzymes with the 2-His-1-carboxylate facial triad, *Dalton Transactions* 3153-3162.

76. Haahr, L. T., Jensen, K. P., Boesen, J., and Christensen, H. E. M. (2010) Experimentally calibrated computational chemistry of tryptophan hydroxylase: trans influence, hydrogen-bonding, and 18-electron rule govern O2-activation, *Journal of Inorganic Biochemistry 104*, 136-145.

77. Olsson, E., Martinez, A., Teigen, K., and Jensen, V. R. (2011) Substrate hydroxylation by the oxido-iron intermediate in aromatic amino acid hydroxylases: A DFT mechanistic study, *European Journal of Inorganic Chemistry 2011*, 2720-2732.

78. Olsson, E., Martinez, A., Teigen, K., and Jensen, V. R. (2011) Formation of the iron-oxo hydroxylating species in the catalytic cycle of aromatic amino acid hydroxylases, *Chemistry 17*, 3746-3758.

79. Eser, B. E., Barr, E. W., Frantom, P. a, Saleh, L., Bollinger, J. M., Krebs, C., and Fitzpatrick, P. F. (2007) Direct spectroscopic evidence for a high-spin Fe(IV) intermediate in tyrosine hydroxylase, *Journal of the American Chemical Society 129*, 11334-11335.

80. Panay, A. J., Lee, M., Krebs, C., Bollinger, J. M., and Fitzpatrick, P. F. (2011) Evidence for a high-spin Fe(IV) species in the catalytic cycle of a bacterial phenylalanine hydroxylase, *Biochemistry 50*, 1928-1933.

81. Lange, S. J., Miyake, H., and Que, L. (1999) Evidence for a nonheme Fe(IV)=O species in the intramolecular hydroxylation of a phenyl moiety, *Journal of the American Chemical Society 121*, 6330-6331.

82. Krebs, C., Galonić Fujimori, D., Walsh, C. T., and Bollinger, J. M. (2007) Non-heme Fe(IV)-oxo intermediates, *Accounts of Chemical Research 40*, 484-492.

83. Francisco, W. A., Tian, G. C., Fitzpatrick, P. F., and Klinman, J. P. (1998) Oxygen-18 kinetic isotope effect studies of the tyrosine hydroxylase reaction: Evidence of rate limiting oxygen activation, *Journal of the American Chemical Society 120*, 4057-4062.

84. Kemsley, J. N., Wasinger, E. C., Datta, S., Mitić, N., Acharya, T., Hedman, B., Caradonna, J. P., Hodgson, K. O., and Solomon, E. I. (2003) Spectroscopic and kinetic studies of PKU-inducing mutants of phenylalanine hydroxylase: Arg158Gln and Glu280Lys, *Journal of the American Chemical Society 125*, 5677-5686.

85. Hillas, P. J., and Fitzpatrick, P. F. (1996) A mechanism for hydroxylation by tyrosine hydroxylase based on partitioning of substituted phenylalanines, *Biochemistry 35*, 6969-6975.

86. Moran, G. R., Derecskei-Kovacs, A., Hillas, P. J., and Fitzpatrick, P. F. (2000) On the catalytic mechanism of tryptophan hydroxylase, *Journal of the American Chemical Society 122*, 4535-4541.

87. Moran, G. R., Phillips, R. S., and Fitzpatrick, P. F. (1999) Influence of steric bulk and electrostatics on the hydroxylation regiospecificity of tryptophan hydroxylase: characterization of methyltryptophans and azatryptophans as substrates, *Biochemistry 38*, 16283-16289.

88. Renson, J., Daly, J., Weissbach, H., Witkop, B., and Udenfriend, S. (1966) Enzymatic conversion of 5-tritiotryptophan to 4-tritio-5-hydroxytryptophan, *Biochemical and Biophysical Research Communications 25*, 504-513.

89. Guroff, G., and Daly, J. (1967) Quantitative studies on the hydroxylation-induced migration of deuterium and tritium during phenylalanine hydroxylation, *Archives of Biochemistry and Biophysics 122*, 212-217.

90. Daly, J., Levitt, M., Guroff, G., and Udenfriend, S. (1968) Isotope studies on the mechanism of action of adrenal tyrosine hydroxylase, *Archives of Biochemistry and Biophysics 126*, 593-598.

91. Bassan, A., Blomberg, M. R. a, and Siegbahn, P. E. M. (2003) Mechanism of aromatic hydroxylation by an activated FeIV=O core in tetrahydrobiopterin-dependent hydroxylases, *Chemistry 9*, 4055-4067.

92. Shiota, Y., and Yoshizawa, K. (2004) QM/MM study of the mononuclear non-heme iron active site of phenylalanine hydroxylase, *Journal of Physical Chemistry B 108*, 17226-17237.

93. Abitas, J.-pierre, Parniaks, M., Kaufmann, S., Abita, J. P., Parniak, M., and Kaufman, S. (1984) The activation of rat liver phenylalanine hydroxylase by limited proteolysis, lysolecithin, and tocopherol phosphate. Changes in conformation and catalytic properties, *The Journal of Biological Chemistry 259*, 14560-14566.

94. Boesen, J. (2007) Cloning, production and chromatographic purification of human tryptophan hydroxylase isoform 1 variants and initial kinetic studies (in Danish), Master thesis. Department of Chemistry, Technical University of Denmark.

95. McPherson, A. (2004) Introduction to protein crystallization, *Methods 34*, 254-265.

96. Nielsen, M. S., Petersen, C. R., Munch, A., Vendelboe, T. V., Boesen, J., Harris, P., and Christensen, H. E. (2008) A simple two step procedure for purification of the catalytic domain of chicken tryptophan hydroxylase 1 in a form suitable for crystallization, *Protein Expression and Purification 57*, 116-126.

97. Yang, X. J., and Kaufman, S. (1994) High-level expression and deletion mutagenesis of human tryptophan hydroxylase, *Proceedings of the National Academy of Sciences of the United States of America 91*, 6659-6663.

98. McKinney, J., Knappskog, P. M., Pereira, J., Ekern, T., Toska, K., Kuitert, B. B., Levine, D., Gronenborn, A. M., Martinez, A., and Haavik, J. (2004) Expression and purification of human tryptophan hydroxylase from Escherichia coli and Pichia pastoris, *Protein Expression and Purification 33*, 185-194.

99. Mockus, S. M., Kumer, S. C., and Vrana, K. E. (1997) A chimeric tyrosine/tryptophan hydroxylase. The tyrosine hydroxylase regulatory domain serves to stabilize enzyme activity, *Journal of Molecular Neuroscience 9*, 35-48.

100. Mockus, S. M., Kumer, S. C., and Vrana, K. E. (1997) Carboxyl terminal deletion analysis of tryptophan hydroxylase, *Biochimica et Biophysica Acta 1342*, 132-140.

101. D'Sa, C. M., Arthur, R. E., and Kuhn, D. M. (1996) Expression and deletion mutagenesis of tryptophan hydroxylase fusion proteins: delineation of the enzyme catalytic core, *Journal of Neurochemistry 67*, 917-926.

102. Winge, I., McKinney, J. a, Knappskog, P. M., and Haavik, J. (2007) Characterization of wild-type and mutant forms of human tryptophan hydroxylase 2, *Journal of Neurochemistry 100*, 1648-1657.

103. Kowlessur, D., and Kaufman, S. (1999) Cloning and expression of recombinant human pineal tryptophan hydroxylase in Escherichia coli: Purification and characterization of the cloned enzyme, *Biochimica et Biophysica Acta 1434*, 317-330.

104. Munch, A. (2004) Affinity ligands for purification of tryptophan hydroxylase. Department of Chemistry, Technical University of Denmark.

105. Waugh, D. S. (2005) Making the most of affinity tags, *Trends in Biotechnology 23*, 316-320.

106. GE Healthcare. (2009) Recombinant Protein Purification Handbook - Principles and methods. GE Healthcare Bio-Sciences AB, Uppsala, Sweden.

107. Waugh, D. S. (2011) An overview of enzymatic reagents for the removal of affinity tags, *Protein Expression and Purification 80*, 283-293.

108. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein, *Protein Science 4*, 2411-2423.

109. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., and Hochstrasser, D. F. (1999) Protein identification and analysis tools in the ExPASy server, *Methods in Molecular Biology 112*, 531-552.

110. BioRad webpage. www.bio-rad.com, 03.04.2012.

111. Waters. (2006) LCT Premier XE Quick Start Guide 71500127003 / Revision A, in *Waters LCT Premier XE Bookshelf*. Waters coperation.

112. Cordingley, M. G., Callahan, P. L., Sardana, V. V., Garsky, V. M., and Colonno, R. J. (1990) Substrate requirements of human rhinovirus 3C protease for peptide cleavage in-vitro, *Journal of Biological Chemistry 265*, 9062-9065.

113. Walker, P. A., Leong, L. E. C., Ng, P. W. P., Tan, S. H., Waller, S., Murphy, D., and Porter, A. G. (1994) Efficient and rapid affinity purification of proteins using recombinant fusion proteases, *Bio-Technology 12*, 601-605.

114. Kaplan, W., Hüsler, P., Klump, H., Erhardt, J., Sluis-Cremer, N., and Dirr, H. (1997) Conformational stability of pGEX-expressed Schistosoma japonicum glutathione S-transferase: a detoxification enzyme and fusion-protein affinity tag, *Protein Science 6*, 399-406.

115. Kapust, R. B., and Waugh, D. S. (1999) Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused, *Protein Science 8*, 1668-1674.

116. Sachdev, D., and Chirgwin, J. M. (1998) Order of fusions between bacterial and mammalian proteins can determine solubility in Escherichia coli, *Biochemical and Biophysical Research Communications 244*, 933-937.

117. Schein, C. H., and Noteborn, M. H. M. (1988) Formation of soluble recombinant proteins in Escherichia coli is favored by lower growth temperatures, *Biotechnology 6*, 291-294.

118. Hernández, H., and Robinson, C. V. (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry, *Nature Protocols 2*, 715-726.

119. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, *Nature Protocols 2*, 2212-2221.

120. Bright, G. R., Fisher, G. W., Rogowska, J., and Taylor, D. L. (1987) Fluorescence ratio imaging microscopy: temporal and spatial measurements of cytoplasmic pH, *The Journal of Cell Biology 104*, 1019-1033.

121. Joh, T. H., Shikimi, T., Pickel, V. M., and Reis, D. J. (1975) Brain tryptophan hydroxylase: purification of, production of antibodies to, and cellular and ultrastructural localization in serotonergic neurons of rat midbrain, *Proceedings of the National Academy of Sciences of the United States of America 72*, 3575-3579.

122. Bergfors, T. (Ed.). (1999) Protein Crystallization: Techniques, Strategies, and Tips - A Laboratory Manual. IUL Biotechnology Series, La Jolla, California.

123. Jancarik, J., and Kim, S.-H. (1991) Sparse matrix sampling. A screening method for crystallization of proteins, *Journal of Applied Crystallography 24*, 409-411.

124. Rupp, B., and Wang, J. (2004) Predictive models for protein crystallization, *Methods 34*, 390-407.

125. Chayen, N. E., and Saridakis, E. (2008) Protein crystallization: from purified protein to diffraction-quality crystal, *Nature Methods 5*, 147-153.

126. Jin, H., Cianchetta, G., Devasagayaraj, A., Gu, K., Marinelli, B., Samala, L., Scott, S., Stouch, T., Tunoori, A., Wang, Y., Zang, Y., Zhang, C., David Kimball, S., Main, A. J., Ding, Z.-M. M., Sun, W., Yang, Q., Yu, X.-Q. Q., Powell, D. R., Wilson, A., Liu, Q., and Shi, Z.-C. C. (2009) Substituted 3-(4-(1,3,5-triazin-2-yl)-phenyl)-2-aminopropanoic acids as novel tryptophan hydroxylase inhibitors, *Bioorganic & Medicinal Chemistry Letters 19*, 5229-5232.

127. Goodwill, K. E., Sabatier, C., and Stevens, R. C. (1998) Crystal structure of tyrosine hydroxylase with bound cofactor analogue and iron at 2.3 A resolution: self-hydroxylation of Phe300 and the pterin-binding site, *Biochemistry 37*, 13437-13445.

128. Erlandsen, H., Fusetti, F., Martinez, A., Hough, E., Flatmark, T., and Stevens, R. C. (1997) Crystal structure of the catalytic domain of human phenylalanine hydroxylase reveals the structural basis for phenylketonuria, *Nature Structural Biology 4*, 995-1000.

129. Erlandsen, H., Flatmark, T., Stevens, R. C., and Hough, E. (1998) Crystallographic analysis of the human phenylalanine hydroxylase catalytic domain with bound catechol inhibitors at 2.0 angstrom resolution, *Biochemistry 37*, 15638-15646.

130. Erlandsen, H., Bjorgo, E., Flatmark, T., and Stevens, R. C. (2000) Crystal structure and site-specific mutagenesis of pterin-bound human phenylalanine hydroxylase, *Biochemistry 39*, 2208-2217.

131. Erlandsen, H., Kim, J. Y., Patch, M. G., Han, A., Volner, A., Abu-Omar, M. M., and Stevens, R. C. (2002) Structural comparison of bacterial and human iron-dependent phenylalanine hydroxylases: Similar fold, different stability and reaction rates, *Journal of Molecular Biology 320*, 645-661.

132. Andersen, O. A., Flatmark, T., and Hough, E. (2002) Crystal structure of the ternary complex of the catalytic domain of human phenylalanine hydroxylase with tetrahydrobiopterin and 3-(2-thienyl)-L-alanine, and its implications for the mechanism of catalysis and substrate activation, *Journal of Molecular Biology 320*, 1095-1108.

133. Erlandsen, H., Pey, A. L., Gamez, A., Perez, B., Desviat, L. R., Aguado, C., Koch, R., Surendran, S., Tyring, S., Matalon, R., Scriver, C. R., Ugarte, M., Martinez, A., and Stevens, R. C. (2004) Correction of kinetic and stability defects by tetrahydrobiopterin in phenylketonuria patients with certain phenylalanine hydroxylase mutations, *Proceedings of the National Academy of Sciences of the United States of America 101*, 16903-16908.

134. Leiros, H. K. S., Pey, A. L., Innselset, M., Moe, E., Leiros, I., Steen, I. H., and Martinez, A. (2007) Structure of phenylalanine hydroxylase from Colwellia psychrerythraea 34H, a monomeric cold active enzyme with local flexibility around the active site and high overall stability, *Journal of Biological Chemistry 282*, 21973-21986.

135. Ursby, T., Mammen, C. B., Cerenius, Y., Svensson, C., Sommarin, B., Fodje, M. N., Kvick, A., Logan, D. T., Als-Nielsen, J., and Thunnissen, M. M. G. M. (2004) The new macromolecular crystallography stations at MAX-lab: The MAD station, *AIP Conference Proceedings 705*, 1241-1246.

136. Robinson, H., Soares, A. S., Becker, M., Sweet, R., and Heroux, A. (2006) Mail-in crystallography program at brookhaven national laboratory's national synchrotron light source, *Acta Crystallographica Section D Biological Crystallography 62*, 1336-1339.

137. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. W. (2011) iMOSFLM: a new graphical interface for diffraction image processing with MOSFLM, *Acta Crystallographica, Section D (Biological Crystallography) D67*, 271-281.

138. Kabsch, W. (2010) XDS, *Acta Crystallographica, Section D (Biological Crystallography) D66*, 125-132.

139. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite, *Acta Crystallographica Section D Biological Crystallography 59*, 1131-1137.

140. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallographica, Section D (Biological Crystallography) D66*, 486-501.

141. Schmit, J. D., and Dill, K. (2012) Growth rates of protein crystals, *Journal of the American Chemical Society 134*, 3934-3937.

142. Vergara, A., Lorber, B., Zagari, A., and Giegé, R. (2002) Physical aspects of protein crystal growth investigated with the Advanced Protein Crystallization Facility in reduced-gravity environments, *Acta Crystallographica Section D Biological Crystallography 59*, 2-15.

143. Brünger, A. T. (1997) Free R value: cross-validation in crystallography, *Methods in Enzymology 277*, 366-396.

144. P-CUBE (Infrastructure for Protein Production Platforms). http://www.p-cube.eu/, 21.05.2012.

145. Gräslund, S., Sagemark, J., Berglund, H., Dahlgren, L.-G., Flores, A., Hammarström, M., Johansson, I., Kotenyova, T., Nilsson, M., Nordlund, P., and Weigelt, J. (2008) The use of systematic N- and C-terminal deletions to promote production and structural studies of recombinant proteins, *Protein Expression and Purification 58*, 210-221.

146. McKinney, J. A., Turel, B., Winge, I., Knappskog, P. M., and Haavik, J. (2009) Functional properties of missense variants of human tryptophan hydroxylase 2, *Human Mutation 30*, 787-794.

147. Daubner, S. C., and Fitzpatrick, P. F. (1999) Site-directed mutants of charged residues in the active site of tyrosine hydroxylase, *Biochemistry 38*, 4448-4454.

148. Daubner, S. C., and Fitzpatrick, P. F. (1998) Mutation to phenylalanine of tyrosine 371 in tyrosine hydroxylase increases the affinity for phenylalanine, *Biochemistry 37*, 16440-16444.

149. Kinzie, S. D., Thevis, M., Ngo, K., Whitelegge, J., Loo, J. a, and Abu-Omar, M. M. (2003) Posttranslational hydroxylation of human phenylalanine hydroxylase is a novel example of enzyme self-repair within the second coordination sphere of catalytic iron, *Journal of the American Chemical Society 125*, 4710-4711.

150. Miranda, F. F., Kolberg, M., Andersson, K. K., Geraldes, C. F. G. C., and Martínez, A. (2005) The active site residue tyrosine 325 influences iron binding and coupling efficiency in human phenylalanine hydroxylase, *Journal of Inorganic Biochemistry 99*, 1320-1328.

151. Li, J., and Fitzpatrick, P. F. (2008) Characterization of metal ligand mutants of phenylalanine hydroxylase: Insights into the plasticity of a 2-histidine-1-carboxylate triad, *Archives of Biochemistry and Biophysics 475*, 164-168.

152. Fitzpatrick, P. F., Ralph, E. C., Ellis, H. R., Willmon, O. J., and Daubner, S. C. (2003) Characterization of metal ligand mutants of tyrosine hydroxylase: Insights into the plasticity of a 2-histidine-1-carboxylate triad, *Biochemistry 42*, 2081-2088.

153. Gibbss, B. S., Wojchowski, D., and Benkovic, S. J. (1993) Expression of rat liver phenylalanine hydroxylase in insect cells and site-directed mutagenesis of putative non-heme, *Journal of Biological Chemistry 268*, 8046-8052.

154. Ramsey, A. J., Daubner, S. C., Ehrlich, J. I., and Fitzpatrick, P. F. (1995) Identification of iron ligands in tyrosine hydroxylase by mutagenesis of conserved histidinyl residues, *Protein Science 4*, 2082-2086.

155. Diebold, A. R., Neidig, M. L., Moran, G. R., Straganz, G. D., and Solomon, E. I. (2010) The three-his triad in Dke1: comparisons to the classical facial triad, *Biochemistry 49*, 6945-6952.

156. McCoy, J. G., Bailey, L. J., Bitto, E., Bingman, C. A., Aceti, D. J., Fox, B. G., and Phillips Jr, G. N. (2006) Structure and mechanism of mouse cysteine dioxygenase, *Proceedings of the National Academy of Sciences of the United States of America 103*, 3084-3089.

157. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data, *Analytical Biochemistry 182*, 319-326.

158. Solomon, E. I., Pavel, E. G., Loeb, K. E., and Campochiaro, C. (1995) Magnetic circular dichroism spectroscopy as a probe of the geometric and electronic structure of non-heine ferrous enzymes, *Coordination Chemistry Reviews 44*, 369-460.

159. Solomon, E. I., and Hanson, M. A. (1999) Bioinorganic spectroscopy, in *Inorganic electronic structure and spectroscopy, Volume II: Applications and Case Studies* (Solomon, E. I., and Lever, A. B. P., Eds.), pp 1-129. John Wiley & Sons, Inc., New York.

160. Neidig, M. L., and Solomon, E. I. (2005) Structure-function correlations in oxygen activating non-heme iron enzymes, *Chemical Communications* 5843-5863.

161. Solomon, E. I., Brunold, T. C., Davis, M. I., Kemsley, J. N., Lee, S. K., Lehnert, N., Neese, F., Skulan, A. J., Yang, Y. S., and Zhou, J. (2000) Geometric and electronic structure/function correlations in non-heme iron enzymes, *Chemical Reviews 100*, 235-350.

162. Pavel, E., Kitajima, N., and Solomon, E. I. (1998) Magnetic circular dichroism spectroscopic studies of mononuclear non-heme ferrous model complexes. Correlation of excited- and ground- state electronic structure with geometry, *Journal of the American Chemical Society 120*, 3949-3962.

163. Bucci, E. The circular dichroism site. http://www.imb-jena.de/ImgLibDoc/cd/tut1b.html, 15.05.2012, *Jena Library of Biological Macromolecules*.

164. Johnson, M. K. (2000) CD and MCD Spectroscopy, in *Physical methods in bioinorganic chemistry - spectroscopy and magnetism* (Que, L., Ed.), pp 233-286. University Science Books, Sausalito, California.

165. Piepho, S. B., and Schatz, P. N. (1983) Group theory in spectroscopy: With applications to magnetic circular dichroism. John Wiley, New York.

166. Vagenende, V., Yap, M. G. S., and Trout, B. L. (2009) Mechanisms of protein stabilization and prevention of protein aggregation by glycerol, *Biochemistry 48*, 11084-11096.

167. Bendtsen, N. T. (2011) Computational studies of tryptophan hydroxylase, Master thesis. DTU Chemistry, Technical University of Denmark.

168. Andersson, K. K., Cox, D. D., Que, L. J. R., Flatmark, T., and Haavik, J. (1988) Resonance raman studies on the blue-green-colored bovine adrenal tyrosine 3-monooxygenase tyrosine hydroxylase. Evidence that the feedback inhibitors adrenaline and noradrenaline coordinate to iron, *Journal of Biological Chemistry 263*, 18621-18626.

169. Nakamura, K., Sugawara, Y., Sawabe, K., Ohashi, A., Tsurui, H., Xiu, Y., Ohtsuji, M., Lin, Q. S., Nishimura, H., Hasegawa, H., and Hirose, S. (2006) Late developmental stage-specific role of tryptophan hydroxylase 1 in brain serotonin levels, *The Journal of Neuroscience 26*, 530-534.

170. Tenner, K., Walther, D. J., and Bader, M. (2007) Structural and functional determinants of human tryptophan hydroxylase 2 activity, *FASEB Journal 21*, A1178.

171. Freyer, M. W., and Lewis, E. A. (2008) Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions, *Methods in Cell Biology 84*, 79-113.

172. MicroCal. The Use of Reducing Agents. User manual.

173. GE Healthcare. (2012) MicroCal webpage on ITC. http://www.microcal.com/technology/itc.asp, 15.05.2012, *What is ITC?*

174. Moran, G. R., and Fitzpatrick, P. F. (1999) A continuous fluorescence assay for tryptophan hydroxylase, *Analytical Biochemistry 266*, 148-152.

175. Deng, H., Callender, R., and Dale, G. E. (2000) A vibrational structure of 7,8-dihydrobiopterin bound to dihydroneopterin aldolase, *Journal of Biological Chemistry 275*, 30139-30143.

176. Davis, M. D., Kaufman, S., and Milstien, S. (1988) The auto-oxidation of tetrahydrobiopterin, *European Journal of Biochemistry 173*, 345-351.

177. Pfleiderer, W. (1987) Chemistry of pterins, in *Unconjugated pterins in neurobiology* (Lovenberg, W., and Levine, R. A., Eds.). Taylor and Francis Ltd, London.

9 Appendix

9.1 Appendix: Protein amino acid sequences

Color code:

XXXX = GST
XXXX = MBP
XXXX = 3CP
XXXX = Linker
XXXX = 3CP recognition sequence and cleavage site (marked with ↔)
XXXX = Regulatory domain of TPH
XXXX = Catalytic domain of TPH
XXXX = Tetramerization domain of TPH

GST-3CP:
MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID
GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKV
DFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFK
KRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLIEGRGIPGPEHEFLNALIR
RNCHIITTDKGEFNLLGIYSNCAVVPTHAEPGDVVDIDGRLVRVLKQQVLTDMNDVDTEV
TVLWLDQNEKFRDIRRFIPEHQQDWHNIHLATNVTKFPMLNVEVGHTVPYGEINLSGNAT
CRLYKYDYPTQPGQCGAVLANTGNIIGIHVGGNGRVGYAAALLRKYFAEEQ

MBP-3CP:
MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDI
IFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNK
DLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIK
DVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSK
VNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDE
ALKDAQTNSSSNNNNNNNNNNLGGIPGPEHEFLNALIRRNCHIITTDKGEFNLLGIYSNC
AVVPTHAEPGDVVDIDGRLVRVLKQQVLTDMNDVDTEVTVLWLDQNEKFRDIRRFIPEHQ
QDWHNIHLATNVTKFPMLNVEVGHTVPYGEINLSGNATCRLYKYDYPTQPGQCGAVLANT
GNIIGIHVGGNGRVGYAAALLRKYFAEEQ

GST-*ch*TPH1:
MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID
GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL
KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC
FKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLEVLFQ↔GPLGSDGMETV
PWFPKKISDLDHCANRVLMYGSELDADHPGFKDNVYRKRRKYFADLAMNYKHGDPIPKVE
FTEEEIKTWGTVFQELNKLYPTHACREYLKNLPLLSKYCGYREDNIPQLEDVSNFLKERT
GFSIRPVAGYLSPRDFLSGLAFRVFHCTQYVRHSSDPFYTPEPDTCHELLGHVPLLAEPS
FAQFSQEIGLASLGASEEAVQKLATCYFFTVEFGLCKQDGQLRVFGAGLLSSISELKHAL
SGHAKVKPFDPKITCKQECLITTFQDVYFVSESFEDAKEKMREFTKTIKRPFGVKYNPYT
RSIQILKD

MBP-*h*TPH1:

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGD
GPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSL
IYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGK
YDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNI
DTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNK
DKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQ
TVDEALKDAQTNSSSNNNNNNNNNNLGLEVLFQ↔GPGSIEDNKENKDHSLERGRASLIFSL
KNEVGGLIKALKIFQEKHVNLLHIESRKSKRRNSEFEIFVDCDINREQLNDIFHLLKSHT
NVLSVNLPDNFTLKEDGMETVPWFPKKISDLDHCANRVLMYGSELDADHPGFKDNVYRKR
RKYFADLAMNYKHGDPIPKVEFTEEEIKTWGTVFQELNKLYPTHACREYLKNLPLLSKYC
GYREDNIPQLEDVSNFLKERTGFSIRPVAGYLSPRDFLSGLAFRVFHCTQYVRHSSDPFY
TPEPDTCHELLGHVPLLAEPSFAQFSQEIGLASLGASEEAVQKLATCYFFTVEFGLCKQD
GQLRVFGAGLLSSISELKHALSGHAKVKPFDPKITCKQECLITTFQDVYFVSESFEDAKE
KMREFTKTIKRPFGVKYNPYTRSIQILKDTKSITSAMNELQHDLDVVSDALAKVSRKPSI

MBP-rc*h*TPH1:

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDI
IFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNK
DLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIK
DVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSK
VNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDE
ALKDAQTNSSSNNNNNNNNNNLGLEVLFQ↔GPGSIEDNKENKDHSLERGRASLIFSLKNEV
GGLIKALKIFQEKHVNLLHIESRKSKRRNSEFEIFVDCDINREQLNDIFHLLKSHTNVLS
VNLPDNFTLKEDGMETVPWFPKKISDLDHCANRVLMYGSELDADHPGFKDNVYRKRRKYF
ADLAMNYKHGDPIPKVEFTEEEIKTWGTVFQELNKLYPTHACREYLKNLPLLSKYCGYRE
DNIPQLEDVSNFLKERTGFSIRPVAGYLSPRDFLSGLAFRVFHCTQYVRHSSDPFYTPEP
DTCHELLGHVPLLAEPSFAQFSQEIGLASLGASEEAVQKLATCYFFTVEFGLCKQDGQLR
VFGAGLLSSISELKHALSGHAKVKPFDPKITCKQECLITTFQDVYFVSESFEDAKEKMRE
FTKTIKRPFGVKYNPYTRSIQILKD

c*h*TPH2:

EELEDVPWFPRKISELDKCSHRVLMYGSELDADHPGFKDNVYRQRRKYFVDVAMGYKYGQ
PIPRVEYTEEETKTWGVVFRELSKLYPTHACREYLKNFPLLTKYCGYREDNVPQLEDVSM
FLKERSGFTVRPVAGYLSPRDFLAGLAYRVFHCTQYIRHGSDPLYTPEPDTCHELLGHV
PLLADPKFAQFSQEIGLASLGASDEDVQKLATCYFFTIEFGLCKQEGQLRAYGAGLLSSI
GELKHALSDKACVKAFDPKTTCLQECLITTFQEAYFVSESFEEAKEKMRDFAKSITRPFS
VYFNPYTQSIEILKD

ct*h*TPH2:

EELEDVPWFPRKISELDKCSHRVLMYGSELDADHPGFKDNVYRQRRKYFVDVAMGYKYGQ
PIPRVEYTEEETKTWGVVFRELSKLYPTHACREYLKNFPLLTKYCGYREDNVPQLEDVSM
FLKERSGFTVRPVAGYLSPRDFLAGLAYRVFHCTQYIRHGSDPLYTPEPDTCHELLGHVP
LLADPKFAQFSQEIGLASLGASDEDVQKLATCYFFTIEFGLCKQEGQLRAYGAGLLSSIG
ELKHALSDKACVKAFDPKTTCLQECLITTFQEAYFVSESFEEAKEKMRDFAKSITRPFSV
YFNPYTQSIEILKDTRSIENVVQDLRSDLNTVCDALNKMNQYLGI

MBP-rc*h*TPH2:

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDI
IFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNK
DLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIK
DVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSK
VNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDE
ALKDAQTNSSSNNNNNNNNNNLGLEVLFQ↔GPGSQPAMMMFSSKYWARRGFSLDSAVPEEH
QLLGSSTLNKPNSGKNDDKGNKGSSKREAATESGKTAVVFSLKNEVGGLVKALRLFQEKR
VNMVHIESRKSRRRSSEVEIFVDCECGKTEFNELIQLLKFQTTIVTLNPPENIWTEEEEL
EDVPWFPRKISELDKCSHRVLMYGSELDADHPGFKDNVYRQRRKYFVDVAMGYKYGQPIP
RVEYTEEETKTWGVVFRELSKLYPTHACREYLKNFPLLTKYCGYREDNVPQLEDVSMFLK
ERSGFTVRPVAGYLSPRDFLAGLAYRVFHCTQYIRHGSDPLYTPEPDTCHELLGHVPLLA
DPKFAQFSQEIGLASLGASDEDVQKLATCYFFTIEFGLCKQEGQLRAYGAGLLSSIGELK
HALSDKACVKAFDPKTTCLQECLITTFQEAYFVSESFEEAKEKMRDFAKSITRPFSVYFN
PYTQSIEILKD

MBP-*h*TPH2:

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDI
IFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNK
DLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIK
DVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSK
VNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDE
ALKDAQTNSSSNNNNNNNNNNLGLEVLFQ↔GPGSQPAMMMFSSKYWARRGFSLDSAVPEEH
QLLGSSTLNKPNSGKNDDKGNKGSSKREAATESGKTAVVFSLKNEVGGLVKALRLFQEKR
VNMVHIESRKSRRRSSEVEIFVDCECGKTEFNELIQLLKFQTTIVTLNPPENIWTEEEEL
EDVPWFPRKISELDKCSHRVLMYGSELDADHPGFKDNVYRQRRKYFVDVAMGYKYGQPIP
RVEYTEEETKTWGVVFRELSKLYPTHACREYLKNFPLLTKYCGYREDNVPQLEDVSMFLK
ERSGFTVRPVAGYLSPRDFLAGLAYRVFHCTQYIRHGSDPLYTPEPDTCHELLGHVPLLA
DPKFAQFSQEIGLASLGASDEDVQKLATCYFFTIEFGLCKQEGQLRAYGAGLLSSIGELK
HALSDKACVKAFDPKTTCLQECLITTFQEAYFVSESFEEAKEKMRDFAKSITRPFSVYFN
PYTQSIEILKDTRSIENVVQDLRSDLNTVCDALNKMNQYLGI

9.2 Appendix: Description of the 1st optimization screen for c*g*TPH1

Description of the 1st optimization screen, which was set up at room temperature and 5 °C. The screen was set up in combiclover trays (4 drop chambers for each well). Conditions in green writing produced crystals at room temperature.
<u>Conditions varying protein concentration</u>: Each dropchamber was added the following: a: nothing, b: 310 µM tryptophan, c: 1600 µM BH_2, d: 155 µM tryptophan+800 µM BH_2. The protein concentration was different in drop chamber a compared to b-d.
<u>Conditions varying PEG type and concentration</u>: The protein concentration was kept constant at 2.8 mg/ml in drop chambers a-d, and each drop chamber was added the following: a: nothing, b: 160 µM tryptophan, c: 160 µM BH_2, d: 160 µM tryptophan+160 µM BH_2.

Protein concentration	Protein concentration	PEG type and concentration	PEG type and concentration
22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 2.8 mg/ml c*g*TPH1 b-d: 2.8 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 2.9 mg/ml c*g*TPH1 b-d: 2.3 mg/ml c*g*TPH1	18 (w/v)% PEG 10000 10 (v/v) % PEG 400 0.2 M imidazole malate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1	45 (v/v) % PEG 400 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1
22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 2.5 mg/ml c*g*TPH1 b-d: 2.0 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 3.0 mg/ml c*g*TPH1 b-d: 2.4 mg/ml c*g*TPH1	25 (v/v) % PEG 400 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1	40 (v/v) % PEG 400 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1
22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 2.6 mg/ml c*g*TPH1 b-d: 2.1 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 3.1 mg/ml c*g*TPH1 b-d: 2.5 mg/ml c*g*TPH1	23 (v/v) % PEG 400 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1	35 (v/v) % PEG 400 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1
22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 2.7 mg/ml c*g*TPH1 b-d: 2.2 mg/ml c*g*TPH1		20 (w/v)% PEG 10000 6 (v/v) % PEG 400 0.2 M imidazole malate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1	30 (v/v) % PEG 400 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1

9.3 Appendix: Overview of the additive screen used in the 2nd optimization screen for c*g*TPH1

Overview of the additive screen MD1-11 from Molecular Dimensions used in the 2nd optimization screen. The screen was set up at room temperature with the original crystallization conditions (22.5 (w/v)% PEG 10000, 0.2 M imidazole malate pH 8.5) and a protein concentration of 2.8 mg/mL. The screen was set up in combiclover trays (4 drop chambers for each well). None of the conditions produced crystals.

Tube #	Compound	Type	Molecular Weight, g/mol	Product concentration [M]	Recommended drop concentration [M]
1	Glycine	Amino acid	75.07	0.5	0.05
2	Glutamine	Amino acid	146.15	0.5	0.05
3	Spermine tetra HCl	Polyamine	348.19	0.1	0.01
4	Spermidine	Polyamine	145.25	0.1	0.01
5	EDTA	Chelating agent	372.24	0.05	0.005
6	Dithiothreitol	Reducing agent	154.25	0.05	0.005
7	Sodium thiocyanate	Chaotropic salt	81.07	0.1	0.01
8	Taurine	Zwitterion	125.15	0.1	0.01
9	Betaine	Linker	135.16	0.1	0.01
10	DMSO	Organic	78.13	100%	5%*
11	β-octylglucoside	Detergent	292.38	5%	0.5%
12	Xylitol	Sugar	152.10	0.2	0.02
13	Potassium iodide	Salt	166.01	1.0	0.1
14	Thymol	Organic/Anti-fungal	150.20	-	2-3 grains†
15	Malonic acid (di-sodium salt)	Di sodium organic salt	148.00	0.5	0.05
16	Ethylenediamine	Amine	60.10	1.0	0.05*
17	Cadmium chloride	Divalent cation	183.31	0.1	0.01
18	Cobalt(II) chloride	Divalent cation	237.93	0.1	0.01
19	Magnesium chloride	Divalent cation	203.31	0.1	0.01
20	Zinc chloride	Divalent cation	136.29	0.1	0.01
21	Dioxane	Organic	88.11	100%	5%*
22	Phenol	Dissociating agent	94.11	0.2	0.01*
23	2-propanol	Organic	60.10	60%	3%*
24	Glycerol	Organic	92.09	20%	2%

*These reagents are volatile and should be pre-mixed with the well solution as well as drop solution, as described in the instructions.

†2 –3 grains of thymol should be added to the well solution and then proceed as if it were a volatile reagent.

Abbreviations: DMSO; Di methyl sulfoxide, **EDTA;** (ethylenediamine tetraacetic acid)

9.4 Appendix: Description of the 3rd optimization screen for c*g*TPH1

Description of the 3rd optimization screen, which was set up at room temperature. The screen was set up in combiclover trays (4 drop chambers for each well). Conditions in green writing produced crystals.
<u>Conditions varying protein concentration</u>: Each drop chamber was added the following: a: nothing, b: 310 µM tryptophan, c: 1600 µM BH_2, d: 155 µM tryptophan+800 µM BH_2. The concentration was different in drop chamber a compared to b-d.
<u>Conditions varying PEG 10000 concentration</u>: The protein concentration was kept constant at 2.8 mg/ml in drop chambers a-d, and each drop chamber was added the following: a: nothing, b: 160 µM tryptophan, c: 160 µM BH_2, d: 160 µM tryptophan+160 µM BH_2.

Protein concentration (imidazole)	Protein concentration (imidazole)	Protein concentration (Tris acetate)	Protein concentration (Tris acetate)	PEG 10000 concentration (Tris acetate)	PEG 10000 concentration (Tris acetate)
22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 2.9 mg/ml c*g*TPH1 b-d: 2.3 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a: 2.8 mg/ml c*g*TPH1 b-d: 2.8 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a: 2.9 mg/ml c*g*TPH1 b-d: 2.3 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1	25 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1
22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 2.5 mg/ml c*g*TPH1 b-d: 2.0 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 3.0 mg/ml c*g*TPH1 b-d: 2.4 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a: 2.5 mg/ml c*g*TPH1 b-d: 2.0 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a: 3.0 mg/ml c*g*TPH1 b-d: 2.4 mg/ml c*g*TPH1	15 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1	27 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1
22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 2.6 mg/ml c*g*TPH1 b-d: 2.1 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 3.1 mg/ml c*g*TPH1 b-d: 2.5 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a: 2.6 mg/ml c*g*TPH1 b-d: 2.1 mg/ml c*g*TPH1	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a: 3.1 mg/ml c*g*TPH1 b-d: 2.5 mg/ml c*g*TPH1	17 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1	29 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1
22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 a: 2.7 mg/ml cgTPH1 b-d: 2.2 mg/ml c*g*TPH1		22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a: 2.7 mg/ml c*g*TPH1 b-d: 2.2 mg/ml c*g*TPH1		20 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1	35 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 a-d: 2.8 mg/ml c*g*TPH1

9.5 Appendix: Description of the 4th optimization screen, part 1 for c*g*TPH1

Description of the 4th optimization screen, part 1, which was set up at room temperature. The screen was set up in combiclover trays (4 drop chambers for each well). Conditions in green writing produced crystals. Only one drop was set up per reservoir.

PEG 10000 concentration	PEG 10000 concentration	Protein concentration	PEG 10000 concentration		
22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2	27 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 312 µM trp	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 312 µM trp	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 312 µM trp
22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	25 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp+700 µM BH_2	24 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	24 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 312 µM trp	24 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.1 mg/ml c*g*TPH1 312 µM trp
35 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	20 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	22.5 (w/v)% PEG 10000 0.2 M imidazole malate pH 8.5 2.1 mg/ml c*g*TPH1 156 µM trp+700 µM BH_2	26 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	26 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 312 µM trp	26 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.3 mg/ml c*g*TPH1 312 µM trp
29 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp		22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp+700 µM BH_2	29 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	29 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 312 µM trp	29 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.3 mg/ml c*g*TPH1 312 µM trp

9.6 Appendix: Describtion of the 4th optimization screen, part 2 for c*g*TPH1

Description of the 4th optimization screen, part 2, which was set up at room temperature. The screen was set up in pre-greased hanging-drop trays. Conditions in green writing produced crystals. Only one drop was set up per reservoir and at conditions assigned 4 µL+4 µL, the drop size was double the usual.

-	-	-	-	-	-
22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	23.4 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	24.3 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	25.2 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	26.1 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	27.0 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp
27.9 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp	25.2 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp 4 µL+4 µL	27.0 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp 4 µL+4 µL	25.2 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp 2 (v/v)% glycerol	27.0 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.8 mg/ml c*g*TPH1 156 µM trp 2 (v/v)% glycerol	27.0 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp 2 (v/v)% glycerol
21.6 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp	23.4 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp	24.3 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp	25.2 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp	19.8 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp
20.7 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp	21.6 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp	23.4 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp 4 µL+4 µL	19.8 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp

9.7 Appendix: Describtion of the 4th optimization screen, part 3 for c*g*TPH1

Description of the 4th optimization screen, part 3, which was set up at room temperature. The screen was set up in sitting drop trays with only one chamber per reservoir. Conditions in green writing produced crystals. At conditions assigned 4 µL+4 µL, the drop size was double the usual.

Trp	Trp	Trp	Trp+BH_2	Trp+BH_2	Trp+BH_2
20.7 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp	21.6 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp 4 µL+4 µL	20.7 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2	21.6 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2 4 µL+4 µL
21.6 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.25 mg/ml c*g*TPH1 156 µM trp	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.25 mg/ml c*g*TPH1 156 µM trp	23.4 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.25 mg/ml c*g*TPH1 156 µM trp 4 µL+4 µL	21.6 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.25 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.25 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2	23.4 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.25 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2 4 µL+4 µL
22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp 4 µL+4 µL	23.4 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp	24.3 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2 4 µL+4 µL	23.4 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2	24.3 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2
21.6 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp 2 (v/v)% glycerol	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp 2 (v/v)% glycerol	23.4 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp 2 (v/v)% glycerol	21.6 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2 2 (v/v)% glycerol	22.5 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.5 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2 2 (v/v)% glycerol	23.4 (w/v)% PEG 10000 0.1 M Tris acetate pH 8.5 2.0 mg/ml c*g*TPH1 156 µM trp+156 µM BH_2 2 (v/v)% glycerol

9.8 Appendix: Composition of structure screen 1

moleculardimensions.com

Structure Screen 1 MD1-01

Tube #	Salt	Buffer	pH	Precipitant
1	0.02 M calcium chloride	0.1 M sodium acetate	4.6	30 % v/v MPD
2	0.2 M ammonium acetate	0.1 M sodium acetate	4.6	30 % w/v PEG 4K
3	0.2 M ammonium sulfate	0.1 M sodium acetate	4.6	25 % w/v PEG 4K
4	None	0.1 M sodium acetate	4.6	2.0 M sodium formate
5	None	0.1 M sodium acetate	4.6	2.0 M ammonium sulfate
6	None	0.1 M sodium acetate	4.6	8 % w/v PEG 4K
7	0.2 M ammonium acetate	0.1 M tri-sodium citrate	5.6	30 % w/v PEG 4K
8	0.2 M ammonium acetate	0.1 M tri-sodium citrate	5.6	30 % v/v MPD
9	None	0.1 M tri-sodium citrate	5.6	20 % v/v 2-propanol, 20%w/v PEG 4K
10	None	0.1 M tri-sodium citrate	5.6	1.0 M ammonium dihydrogen phosphate
11	0.2 M calcium chloride	0.1 M sodium acetate	4.6	20 % v/v 2-propanol
12	None	0.1 M sodium cacodylate	6.5	1.4 M sodium acetate
13	0.2 M tri-sodium citrate	0.1 M sodium cacodylate	6.5	30 % v/v 2-propanol
14	0.2 M ammonium sulfate	0.1 M sodium cacodylate	6.5	30 % w/v PEG 8K
15	0.2 M magnesium acetate	0.1 M sodium cacodylate	6.5	20 % w/v PEG 8K
16	0.2 M magnesium acetate	0.1 M sodium cacodylate	6.5	30 % v/v MPD
17	None	0.1 M imidazole	6.5	1.0 M sodium acetate
18	0.2 M sodium acetate	0.1 M sodium cacodylate	6.5	30 % w/v PEG 8K
19	0.2 M zinc acetate	0.1 M sodium cacodylate	6.5	18 % w/v PEG 8K
20	0.2 M calcium acetate	0.1 M sodium cacodylate	6.5	18 % w/v PEG 8K
21	0.2 M tri-sodium citrate	0.1 M Na HEPES	7.5	30 % v/v MPD
22	0.2 M magnesium chloride	0.1 M Na HEPES	7.5	30 % v/ v 2-propanol
23	0.2 M calcium chloride	0.1 M Na HEPES	7.5	28 % v/v PEG 400
24	0.2 M magnesium chloride	0.1 M Na HEPES	7.5	30 % v/v PEG 400
25	0.2 M tri-sodium citrate	0.1 M Na HEPES	7.5	20 % v/v 2-propanol
26	None	0.1 M Na HEPES	7.5	0.8 M K/Na tartrate
27	None	0.1 M Na HEPES	7.5	1.5 M lithium sulfate
28	None	0.1 M Na HEPES	7.5	0.8 M sodium dihydrogen phosphate/ 0.8 M K dihydrogen phosphate
29	None	0.1 M Na HEPES	7.5	1.4 M tri-sodium citrate
30	None	0.1 M Na HEPES	7.5	2 % v/v PEG 400, 2.0 M ammonium sulfate
31	None	0.1 M Na HEPES	7.5	10 % v/v 2-propanol, 20% w/v PEG 4K
32	None	0.1 M Tris	8.5	2.0 M ammonium sulfate
33	0.2 M magnesium chloride	0.1 M Tris	8.5	30 % w/v PEG 4K
34	0.2 M tri-sodium citrate	0.1 M Tris	8.5	30 % v/v PEG 400
35	0.2 M lithium sulfate	0.1 M Tris	8.5	30 % w/v PEG 4K
36	0.2 M ammonium acetate	0.1 M Tris	8.5	30 % v/v 2-Propanol
37	0.2 M sodium acetate	0.1 M Tris	8.5	30 % w/v PEG 4K
38	None	0.1 M Tris	8.5	8 % w/v PEG 8K
39	None	0.1 M Tris	8.5	2.0 M ammonium dihydrogen phosphate
40	None	None	-	0.4 M K/Na Tartrate
41	None	None	-	0.4 M ammonium dihydrogen phosphate
42	0.2 M ammonium sulfate	None	-	30 % w/v PEG 8K
43	0.2 M ammonium sulfate	None	-	30 % w/v PEG 4K
44	None	None	-	2.0 M ammonium sulfate
45	None	None	-	4.0 M sodium formate
46	0.05 M potassium dihydrogen phosphate	None	-	20 % w/v PEG 8K
47	None	None	-	30 % w/v PEG 1.5K
48	None	None	-	0.2 M magnesium formate
49	1.0 M lithium sulfate	None	-	2 % w/v PEG 8K
50	0.5 M lithium sulfate	None	-	15 % w/v PEG 8K

Abbreviations: Na HEPES; 2-(4-(2-Hydroxyethyl)-1-piperazinyl)ethanesulfonic Acid Sodium Salt, **MPD**; 2,4-methyl pentanediol, **PEG**; Polyethylene glycol (1.5K, 4K and 8K correspond to the molecular weight, in thousands of Daltons, of PEG), **Tris**; 2-Amino-2-(hydroxymethyl)propane-1,3-diol.

Manufacturer's safety data sheets are available upon request.

3